TAMPERE UNIVERSITY OF TECHNOLOGY

Department of Chemistry and Bioengineering

Laboratory of Environmental Engineering and Biotechnology

SANJEEV BISTA

CLONING AND CHARACTERIZATION OF A NOVEL THERMOSTABLE ENDOGLUCANASE FROM *CALDICELLULOSIRUPTOR BESCII* IN HETERLOLOGOUS HOST *ESCHERICHIA COLI*

Masters of Science Thesis

Examiner: Prof. Matti Karp and
MSc Alessandro Cirrana
Examiner and topic approved in
Faculty council meeting on 17th
August 2011

TAMPERE UNIVERSITY OF TECHNOLOGY
Department of Chemistry and Bioengineering/Laboratory of Environmental Engineering and Biotechnology
SANJEEV BISTA: Cloning and characterization of a novel thermostable endoglucanase from *Caldicellulosiruptor bescii* in heterologous host *Escherichia coli*
Master of Science Thesis: 63 pages, 5 Appendix pages
September 2011
Major: Biotechnology
Examiner: Professor Matti Karp and MSc Alessandro Cirrana
Keywords: Biofuel, lignocellulose, cellulase, endoglucanase, thermophilic bacteria

---

Abstract

Caldicellulosiruptor bescii DSM 6725 is a gram positive and asporogenic bacterium that utilizes various polysaccharides and grows efficiently on untreated plant biomass at an optimum temperature of ≈ 80 °C. In this study, cellulase genes coding endo-l, 4-β-D-glucanase and exo-1, 4-β-D-glucanase (cellobiohydrolase) were cloned from the genomic DNA of saccharolytic thermophilic anaerobic bacteria *C.bescii* and expressed in *Escherichia coli* BL21. The endoglucanase gene contains an ORF of 2,268 bp encoding a protein of 755 amino acid residues, with a calculated molecular weight of 82.154 kDA. It carries a typical prokaryotic signal peptide of 30 amino acid residues. The amino acid sequence alignment of this typical endoglucanase showed 71% homology to cel5A (endoglucanase) from *Thermoanaerobacter tengcongensis* MB4 and 65% homology to endoglucanase from *Caldicellulosiruptor saccharolyticus*. Residues Glu187 and Glu289 were identified as key catalytic amino acids by sequence alignment. The apparent molecular mass of the endoglucanase protein when expressed in *E.coli* is found approximately 75 kDa, and the highest CMCase activity was observed in intracellular space with comparison to extracellular space.

The optimum pH and temperature for the enzyme activity of the endoglucanase were 5 and 70 °C, respectively. However, enzyme activity was observed over a broad range of pH values and temperatures. The expressed and purified endoglucanase retained over 90% of its original activity after incubation at 70 °C for 24 hours. This suggests that the endoglucanase from *C.bescii* is thermostable and active at different pH.

Preface

This thesis work was conducted at Department of Chemistry and Bioengineering, Tampere University of Technology, Tampere, Finland. I am extremely thankful to my supervisor, Researcher Alessandro Cirrana and Professor Matti Karp for their encouragement and valuable guidance during the thesis study. This thesis could not have been completed without their support and encouragement.

I would like to gratefully acknowledge all of my coworkers in the laboratory and the staff in the Department of Chemistry and Bioengineering for their assistance and advice. Especially I would like to thank researchers Bobin Abraham and Rahul Krishnan Mangayil for their support during my studies and doing the proof reading of my report. Finally, I would like to thank my family for their support and love during my study.

23rd August 2011 Tampere

SANJVEEV BISTA

Table of Contents

ABBREVIATIONS

GH	Glycoside Hydrolase
E.coli	*Escherichia coli*
CO_2	Carbon dioxide
H_2O	Water
$C_6H_{12}O_6$	Glucose molecule
Mn	Manganese
EC number	Enzyme classification number
CBD	Carbohydrate binding domain
CAZy	Carbohydrate active enzyme database
CBM	Carbohydrate binding domains
T_m	Transition temperature
CMC	Carboxymethyl cellulose
DNS	Dinitrosalicyclic acid
NaCl	Sodium Chloride
dNTPS	Deoxynucleoside triphosphates
PCR	Polymerase chain reaction
Na-K	Sodium-potassium
DNA	Deoxyribonucleic acid
TEMED	Tetramethylethylenediamine
IPTG	Isopropyl β-D-thiogalactosidase
SDS-PAGE	Sodiumdodecylsulphatepolyacrylamide gel electrophoresis
DSMZ	German Collection of Microorganisms and Cell Cultures
N_2-CO_2	Nitrogen- Carbon dioxide
RBS	Ribosomal binding site
Lac PO	Lactose promoter
LB	Luria-Bertani broth
LA	Luria-Bertani agar
Tris-HCl	Trisaminomethane hydrochloride
K_m	Michael's Menten constant
V_{max}	Maximum velocity

1 Introduction

The dominance of fossil fuels in energy supply started with development of automobile or diesel engine a century ago. Vegetable oil presently known as "biodiesel" were used to run early diesel engine, however shift in feedstock to gasoline for diesel engine started with gasoline production in 1920s due to cheap price, availability and its unique properties (Hill et al. 2009). Till then, world's energy demand solely relies on fossil fuels. However, continuous rise in fuel price, increasing energy consumption, realization of finite and depleting oil source and green house gas emission have shifted the focus towards searching sustainable and renewable energy sources (Ajanovic 2011; Khraisheh et al. 2010).

Apart from other energy alternatives like wind, solar, nuclear and hydro technologies microbe-assisted biofuel production has been an attractive alternative for current petroleum based fuels due to sustainability and reduction in green house gas emission. Liquid biofuels like ethanol and biodiesel can be used as modern transportation fuel with little change in current technologies (Carere et al. 2008). Depending upon biomass or feedstock used for production, biofuels fall in three categories and are classified as first, second and third generation. At present, first generation biofuel is produced using food crops like maize, wheat and sugarcane as feedstock and use of this feedstock for energy generation has created fuel versus food debate. Second generation biofuel are produced using lignocellulose biomass also called as non-edible biomass and overcomes the major drawback of first generation biofuel since it does not compete with food crops for feedstock (Ruane et al. 2010). Lipids and fatty acids produced from microbes like algae and bacteria are categorized as third generation biofuels. The major advantage of these microbes is that they grow in liquid medium and do not compete with food and energy crops for arable land (Rubin 2008).

Lignocellulosic biomass is the most abundant polysaccharide on earth, comprising of cellulose, hemicellulose and lignin bonded to each other by covalent and hydrogen bonds in plant cell wall. Lignocellulosic biomasses are often called cellulosic biomass as cellulose covers around 50 % of biomass. Cellulose is a homopolysaccharide of β-D-glucose residues linked together by β-1, 4-glycosidic bond making it highly resistant to enzymatic and chemical hydrolysis (Chandel et al. 2011; Dashtban et al. 2009).

Several prokaryotic and eukaryotic microbes like fungi and bacteria capable of producing cellulolytic enzymes have been studied and identified. Cellulolytic enzymes are classified in three main classes and are endo-acting (endoglucanase), exo-acting (exoglucanase or cellobiohydrolases) and β-glucosidase which all act in synergistic manner

for cellulose hydrolysis (Dashtban et al. 2009; Lynd et al. 2002). Most of cellulase belongs to group "*glycoside hydrolases* (GH) family". Bacterial cellulases are secreted either as individual enzyme polypeptide or multienzyme complex called as "cellulosome".

Till date, cellulase activity has been reported from various bacterial species from different ecological niche. Cellulolytic enzymes from thermophilic bacteria are thermostable, highly specific and potential for industrial application and biorefinery. Thermophilic anaerobes like *Caldicellulosiruptor* species has been reported producing large range of extracellular cellulolytic and hemicellulolytic enzymes as individual polypeptide (Blumer-Schuette et al. 2010)

1.1 Significance and objective

Cellulolytic enzymes are used in industries for color extraction of juice, in detergents, bio-stoning of jeans and pretreatment of cellulosic biomass. The enzyme activity and thermostability of cellulase plays a crucial role in overall bioprocess cost and function (Turner et al. 2007). Thermostable enzymes are found to have high enzyme activity, catalyzing reaction at elevated temperature and remaining stable for prolonged time. At present, application of cellulase in biorefining area has been research of interest for generation of biofuels (Viikari et al. 2007; Wang et al. 2010). Celluloytic enzyme have been isolated from various thermophilic bacteria and C*lostridium thermocellum* species are the most studied which produce cellulase in multienzyme complex (VanFossen et al. 2008). Investigations on thermostable cellulase from *Caldicellulosiruptor sp.* have academic and industrial potential due to high optimum growth temperature, extracellular enzyme pathway, and recent completion of genome sequencing. Unlike most thermophilic bacteria, *Caldicellulosiruptor species* is devoid of multienzyme complex (Viikari et al. 2007).

The main objective of study was:

- Isolation of endoglucanase and exoglucanase gene from *Caldicellulosiruptor bescii* followed by cloning and expression in *Escherichia coli* BL21.
- Further, study proceeds in screening, purification and characterization of cloned cellulase genes.

2 Theoretical Background

2.1 History of fossil fuels

Fossil fuels have been the dominance of energy supply since a century ago (Hill et al. 2009). It started with invention of diesel engine in 1892 by Rudolf Diesel. Vegetable oils were used to run early diesel engines. With the start of gasoline production in 1920s, fuel for early diesel engine shifted to petroleum distillates refined from crude oil during gasoline production. Shift in feedstock was mainly due to cheap price, availability and its unique properties like less viscosity and lighter than vegetable oil. Since then, use of vegetable oil currently known as biofuel was sidelined for decades (Schmidt 2007).

The global oil crises in 1970's and hike of oil price by Arab countries led to four times increment of gas and diesel price. The oil crisis dragged to realization of finite oil resources and supply of the world (Demain et al. 2005). This incident propelled researchers to find an economic way of biofuels production as an alternative energy source (Hill et al. 2009; Timilsina et al. 2011). Ethanol was considered as an alternative to fossil fuel due to clean burning, high octane rating, and low carbon dioxide emission (Demain et al. 2005). Countries like Brazil and USA speed up their plant (biomass) derived ethanol production which was also followed by China, Kenya and Zimbabwe. Later, with slow downfall of crude oil price, incentive for biofuel generation was postponed in most of the countries except Brazil (Timilsina et al. 2011).

In present scenario fossil fuels solely stands as major source of energy. Use of fossil fuel as energy source is becoming a reason of concerns due to non-sustainable fuel reserves, high cost, environmental issues and import dependency from politically unstable countries (Ajanovic 2011; Khraisheh et al. 2010). Reduction in dependency on fossil fuel can be achieved only by multiple approaches like solar, nuclear, hydrogen, wind and biofuels (Patil et al. 2008). Biofuels can be a suitable alternative since liquid biofuels like ethanol and biodiesel can be used directly for transportation fuel with minimum change in current technologies (Carere et al. 2008).

2.2 Biofuels

Biofuels are defined as biomass derived fuel which can have liquid, or gaseous form like bioethanol, biodiesel, biobutanol, vegetable oils, biomethanol, biogas, biohydrogen and biomethane (Demirbas 2009a; Patil et al. 2008). Biofuels can be categorized as primary and secondary biofuels (Figure 1). Primary biofuels comprise of natural and unprocessed biomass such as fuel-wood, woodchips, and pellets used for general household purposes. On the other hand, secondary biofuels are produced by processing of biomass. Biofuel offers various benefits in economic and environmental impacts to provide energy security to future generation of mankind. Importance and benefits of biofuel in contrast to fossil fuel are listed in Table 1 (Nigam et al. 2011).

Table 1. Importance and benefits of biofuels (Demirbas 2009b).

Economic Impacts	Sustainability Fuel diversity Increased number of rural manufacturing jobs Increased income taxes Increased investment in plant and equipments Agricultural development Reducing the dependency on improved petroleum
Environmental Impacts	Greenhouse gas reduction Reducing air pollution Biodegradability Higher combustion efficiency Improved land and water use Carbon sequestering
Energy Security	Domestic targets Supply reliability Reducing use of fossil fuel Ready availability Domestic distribution Renewability

2.2.1 First generation biofuels

First generation biofuels are normally produced from stored carbohydrate material in plants. These stored carbohydrates are present in seeds, fruits and grains that covers only small fraction of plant biomass as mentioned in Figure 1. Production of first generation biofuels using plant biomass requires relatively simple processing technique; however, the yield from small portion of biomass is insufficient and not considered as sustainable alternative fuel source (Henry 2010; Ruane et al. 2010). In most cases, food crops or only parts of plant biomass are used as substrate for production of first generation biofuels. Ethanol is one of the most well known first generation biofuel which is produced from sugarcane in Brazil and corn in USA. Biodiesel produced from vegetable oils by transesterification process is another first generation biofuel produced in industrial scale (Sims et al. 2010). In both, bioethanol and biodiesel production only certain part of plant biomass is utilized as a result of which issues like food-versus-fuel debate and rise in global food price has arise (Naik et al. 2010; Nigam et al. 2011).

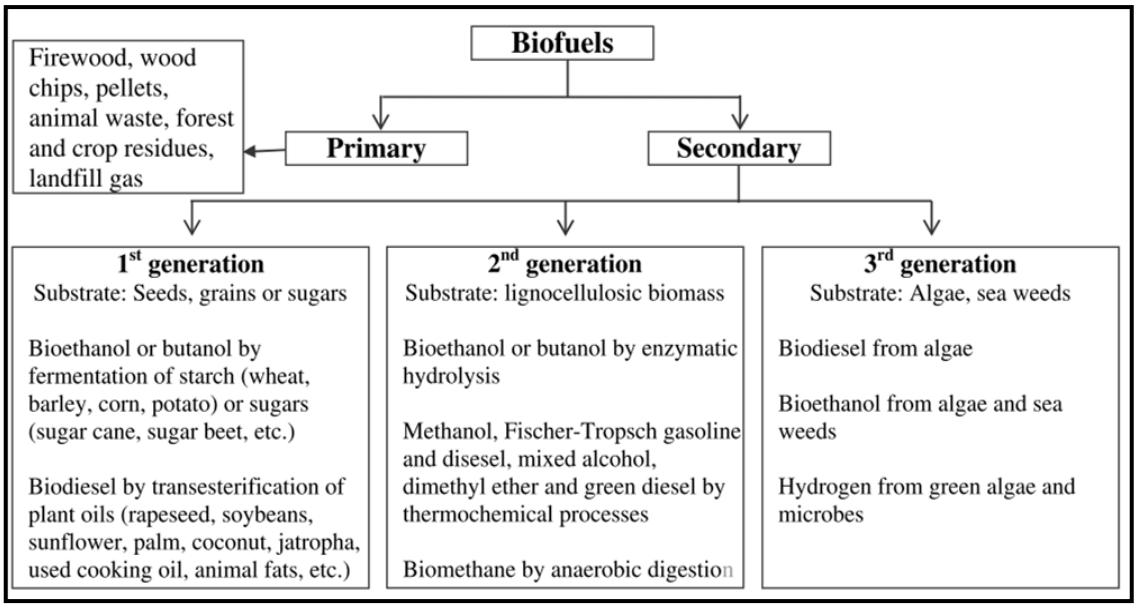


Figure 1. Classification of biofuels (Nigam et al. 2011).

2.2.2 Second generation biofuels

Production of second generation biofuel is carried out by biological or thermochemical means using non-edible plant biomass as substrate. These substrates are collectively known as lignocellulosic biomass which includes crop waste (whole plant biomass), forestry residues and municipal waste materials (Weber et al. 2010). The major advantage of second generation biofuel over first generation is the substrate availability and limited food versus fuel competition for substrates encountered in first generation biofuels (Naik et al. 2010). Two main sources of the lignocellulosic biomass are products from agriculture residues and plant feedstock grown as a substrate for biofuel production (Ruane et al. 2010).

Second generation biofuels like ethanol and butanol are mostly produced through thermochemical means, while the production through biochemical processes is still limited. Currently, fossil fuel is used to produce many second generation thermochemical fuels. Methanol, Fischer-Tropsch liquids (FTL) and dimethyl ether (DME) are few examples of thermochemically produced second generation biofuels. Beside these, pyrolysis oils, often called as unrefined oils are generated thermochemically however, it requires additional refining before using in modern engines (Nigam et al. 2011).

2.2.3 Third generation biofuels

Biofuel derived from microbes and microalgae are called third generation biofuels which includes both liquid and gaseous fuels like hydrogen, methane, ethanol and diesel (Figure 2). They are considered as an alternative energy source for biofuel generation with potential to solve issues associated with first and second generation fuels. Generation of biodiesel from yeast, fungi, and microalgae is feasible since studies have found that microbes are capable of producing and storing large amounts of fatty acid in biomass (Nigam et al. 2011). The characteristic of biodiesel is almost similar to diesel derived from fossil fuel and it can be directly employed in modern engines. The main features of biodiesel are (Ahmad et al. 2011):

- Renewable fuel.
- Highly biodegradable with minimum toxicity.

Microalgae are considered as promising feedstock for better yield of biofuels due to high photosynthetic efficiency leading to fast regeneration time, biomass formation, and higher content of lipids in the cells (20-50% by weight of dry biomass). Unlike other feedstock, it does not compete with food crops for arable land and helps in removal of atmospheric carbon dioxide as well (Ahmad et al. 2011; Gong et al. 2011; Patil et al. 2008).

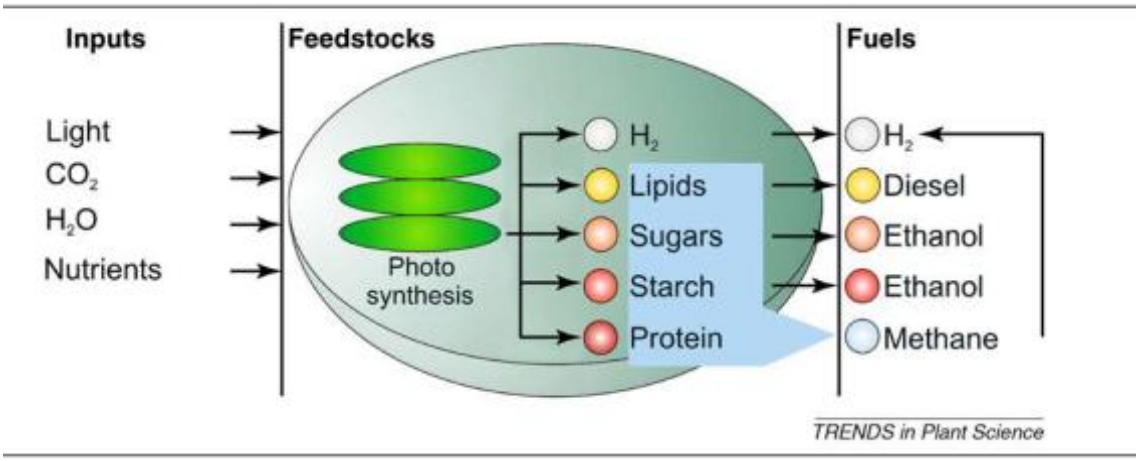


Figure 2. Third generation biofuel production from microalgae (Stephens et al. 2010).

2.3 Feedstock for biofuel generation

There exist various substrate alternatives for biofuel production. However, question arises whether these substrate are sustainable for biofuel production. One of the long- term goals for sustainable production of biofuels is selection of proper feedstock. Renewable sources should be utilized for biofuel production since natural bioresources are available worldwide in comparison to fossil fuels (Nigam et al. 2011).

Bioconversion of lignocellulose into fermentable sugars is considered as potential alternative for generation of sustainable biofuel production (Dellomonaco et al. 2010). Lignocellulose refers to structure of biomass and it comprises of agriculture residues (straws, hulls, stems, stalks etc.), dedicated energy crops (switch grass, and Bermuda grass), municipal solid waste (food and kitchen waste, paper card board, yard trash and wood products etc.), deciduous and coniferous woods, and waste from paper and pulp industry (Badal 2004; Chandel et al. 2011). The production of lignocellulose biomass (plant biomass) is carried out by photosynthesis reaction where light energy from sun in converted into chemical energy and stored as carbohydrates in plant biomass as shown in equation (1) (Zhang 2008). The estimated worldwide annual production of lignocellulose biomass is ≈ 200× 10^9 tons (Chandel et al. 2011; Zhang 2008). Figure 3 shows the source of feedstock for lignocellulose, different sugar present in it and utilization of these sugars for generation of biofuel by various microorganisms.

$$6CO_2 + 6H_2O \rightarrow C_6H_{12}O_6(carbohydrate) + 6O_2 \qquad (1)$$

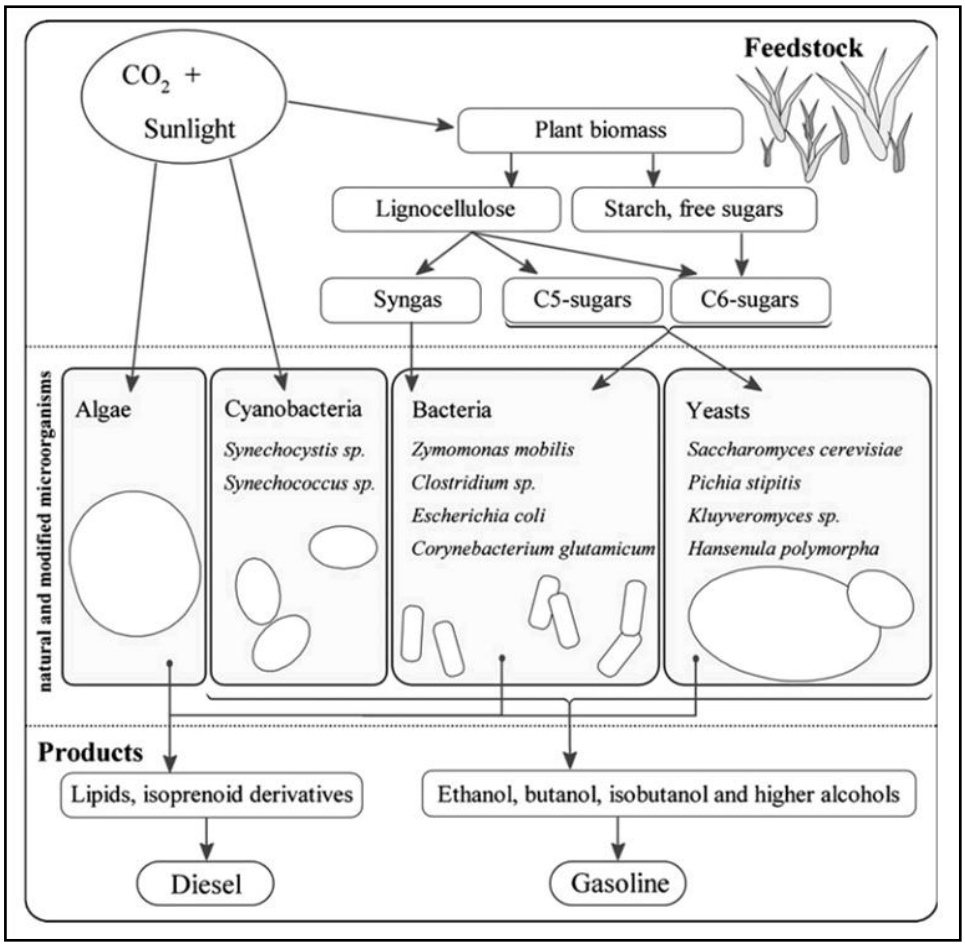


Figure 3. Schematic representation of microbial biofuel production from different feedstock (Weber et al. 2010).

Lignocellulose biomass is the most abundant and renewable source of polysaccharides. It is a complex mixture of three polymers: cellulose, hemicellulose and lignin. Cellulose and hemicellulose are tightly bound to lignin by strong hydrogen bond and few covalent bonds (Lee 1997). These three sugar polymers are associated intimately in order to provide structural framework to the plant cell wall. It is estimated that lignocellulose is composed of up to 75% carbohydrate and around 70 % of plant biomass contains five or six carbon sugars in lignocellulose biomass (Jørgensen et al. 2007; Maki et al. 2009).

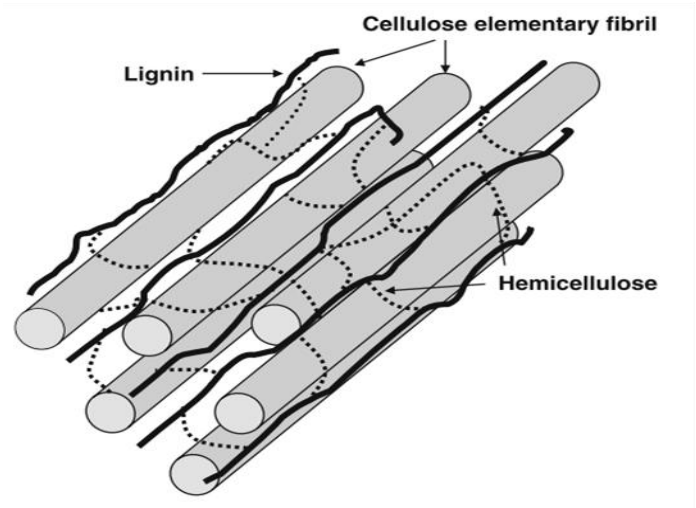


Figure 4. Structure and organization of different sugars in lignocellulose (Vidal et al. 2011).

2.4 Constituents of lignocellulose biomass

The major components of lignocellulose biomass are cellulose (35-50%), hemicellulose (20-35%) and lignin (10-25%) as shown in Table 2. Beside these components, small amount of ash, pectin and proteins occurs in lignocellulose residues (Chandel et al. 2011; Dashtban et al. 2009). The proportion of cellulose, hemicellulose and lignin varies depending upon source of lignocellulose biomass like plant species, age and growth condition.

Table 2. Composition of different sugars in lignocellulose biomass (Chandel et al. 2011).

Biomass type	**Cellulose (%)**	**Hemicellulose (%)**	**Lignin (%)**
Hardwood stems	40–55	22–40	18–25
Softwood stems	45–50	25–35	25–35
Grasses	25–40	25–50	10–30
Waste paper from pulp industry	60 - 70	10- 20	5-10

2.4.1 Lignin

Lignin the third major component of lignocellulosic biomass is formed by complex molecules containing phenyl propane units arranged in three dimensional structures by ether bonds which makes it recalcitrant to biodegradation (Taherzadeh et al. 2008). The major components in lignin are three aromatic alcohols: *p*-coumaryl acohol (*p*-hydroxyphenyl propanol), coniferyl alcohol (guaiacyl propanol) and sinapyl alcohol (syringyl propanol) represented in Figure 5 (Jeffries 1994).

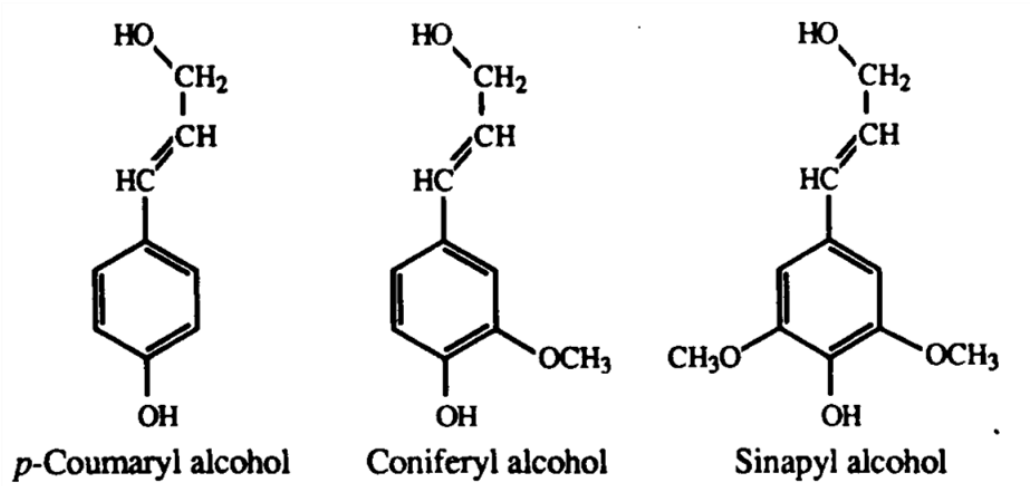


Figure 5. Monomers of lignin polymer (Jeffries 1994).

Lignin acts as glue that helps to bind cellulose and hemicellulose together making it the most recalcitrant component of plant cell wall and resistant to chemical and enzymatic hydrolysis (Taherzadeh et al. 2008). It acts as a defense mechanism and protects plant cells from microbial attack. However, few microorganisms like fungi and bacteria can degrade it (Badal 2004). Softwood plants contain highest amount of lignin than hardwood and herbaceous plants, such as grasses, have the lowest lignin content (Jørgensen et al. 2007). Biodegradation of lignin has been reported from several fungal enzymes like lignin peroxidase, Manganese (Mn)-dependent peroxidase and laccase (mono-phenol oxidase). Among the most widely studied lignin degrading microorganisms, white rot fungi like *Coriolus versicolor*, *Phanerochaete chrysosporium*, and *Trametes versicolor* have been found to be the most efficient ones (Dashtban et al. 2009). Biodegradation of lignin by these enzymes depends upon strain of microbe, accessibility of enzyme and culture conditions (Lee 1997).

2.4.2 Hemicellulose

Hemicellulose is a structural polysaccharide present in plant cell wall along with cellulose and lignin. Hemicellulose components in plant cell wall are mixture of pentoses, hexoses and sugar acids. It is highly branched or linear heteropolysaccharide polymer composed of various sugar residues. These sugar residues can be pentoses (D-xylose, D-arabinose), hexoses (D-mannose, D-glucose, D-galactose) and uronic acids (D-gulcoronic acid and 4-O-methyl-d-glucoronic acid) (Badal 2004; Jeffries 1994; Weber et al. 2010). In plant cell, hemicellulose are synthesized in Golgi apparatus and excreted into cell wall (Minic et al. 2006).

Classification of hemicellulose is carried out on basis of main sugar units present in polymer backbone. The dominant sugars present in hemicellulose are xylan and mannan. Xylan is the dominant hemicellulose sugar in hardwood plant species whereas mannan in softwood species (Kumar et al. 2008). Mannan is an important component of hemicellulose family and can be further classified in four groups: linear mannan, glucomannan, galactomannan and galactoglucomannan. All four polysaccharide has a β-1, 4-linked backbone containing mannose or a combination of glucose and mannose residue. In some cases, side chain of α-1, 6-linked galactose residue can substitute the mannan backbone as shown in Figure 6 (Moreira et al. 2008).

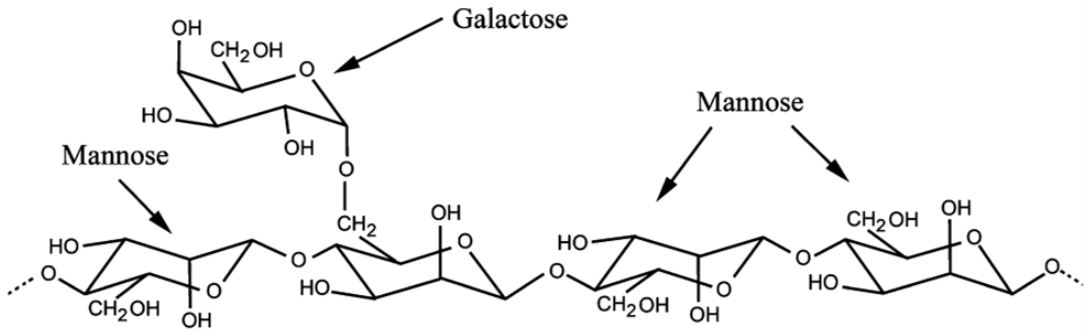


Figure 6. Structure of mannan on hemicellulose backbone (Moreira et al. 2008).

Xylan is the main pentose sugar present in hemicellulose backbone. The structure of xylan consists of β-(1-4)-linked D-xylopyranosyl units with a varying degree of substitution with L-arabinofuranose, glucoronic acid, 4-O- methylglucoronic acid, and acetyl side groups (Bastawde 1992). Xylan is abundant in hardwood species and agricultural waste products like straw and corn stover (Jørgensen et al. 2007). Structure of xylan and hydrolysis site of xylanase enzyme in xylan backbone is shown in Figure 7 (DeBoy et al. 2008).

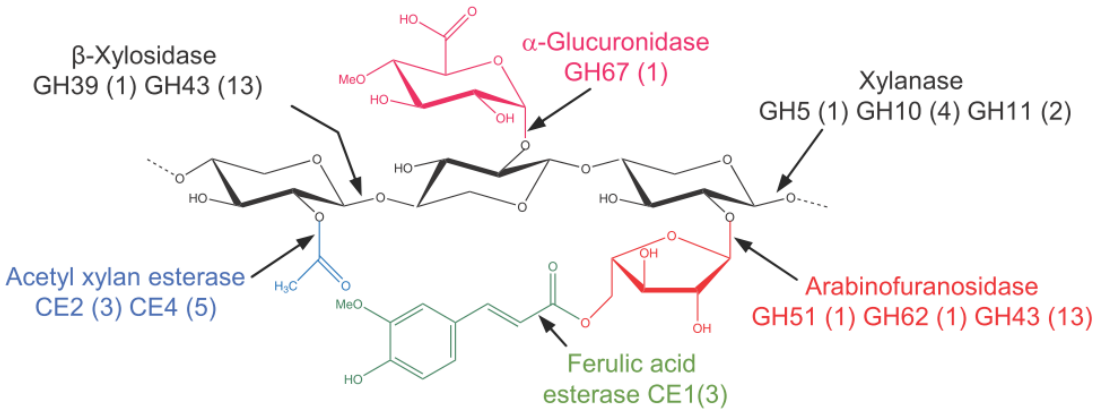


Figure 7. Structure of xylan and site of action of xylanase enzymes in xylanase backbone.

2.4.3 Cellulose

Lignocellulose biomass composition can vary in terms of cellulose, hemicellulose and lignin depending on the source, but cellulose is always predominant and covers around 45 percentage of biomass composition. Cellulose is considered as the most abundant biopolymer on earth synthesized by converting CO_2 and H_2O through photosynthesis with an estimated annual production of 7.5×10^{10} tons (Cao et al. 2002; Carere et al. 2008). Cellulose is a major polymer present in plant cell wall providing structural support and also present in bacteria, fungi and algae. In plant cell, it is synthesized at plasma membrane level and then deposited into cell wall (Agbor et al. 2011; Minic et al. 2006).

Structurally, cellulose is composed of D- glucose monomer linked together with β-1, 4 glycosidic bonds as shown in Figure 8 (Badal 2004; Jørgensen et al. 2007; Kumar et al. 2008). In comparison to other glucan polymers, the repeating unit in cellulose is disaccharide cellobiose instead of glucose. This cellobiose molecule is formed by 180° rotation of β-1, 4 glycosidic bonds between glucose molecules. The degree of polymerization in cellulose polymer can reach length greater than 25,000 glucose residues. The 180° rotation of β-1, 4 glycosidic bonds in linear chain of cellulose results in formation of large amount of both intra and intermolecular hydrogen bonds due to the exposure of OH groups. These hydrogen bonds result in formation of crystalline structure in cellulose, making it insoluble in most solvents and also resistant to microbial enzymatic hydrolysis (Agbor et al. 2011; Carere et al. 2008; Gowen et al. 2010; Jørgensen et al. 2007).

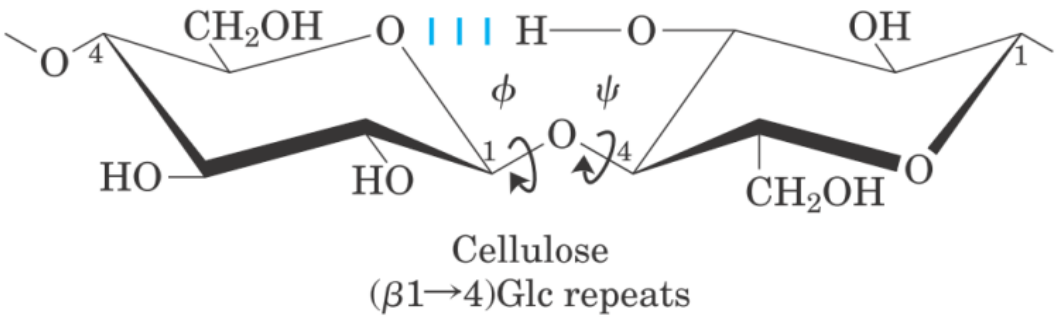


Figure 8. Glycosidic bonds conformation in cellulose backbone (Lehninger et al. 2008).

Crystalline structure of cellulose represents its unique feature in the polysaccharide world. In nature it is biosynthesized as an individual molecule (linear chain of glucosyl molecules) which undergoes self assembly process. Around 30 individual molecules get assembled into larger units known as elementary fibrils (protofibrils) which are packed to form larger units known as microfibrils. Finally, the cellulose fibers are formed by assembly of microfibrils. Occurrence of cellulose fibers in nature is not purely crystalline, although cellulose forms distinct crystalline structure. The shift in degree of crystallinity in cellulose fiber is variable and this results in formation of purely amorphous structure from purely crystalline. Apart from crystalline and amorphous structure, cellulose fibers contains different structural irregularities like kinks or twists of micro fibril, void such as surface micropores, large pits and capillaries apart from crystalline and amorphous structure (Lynd et al. 2002).

3 Biochemistry of hemicellulase and cellulase

The collective groups of enzymes acting on hemicellulose backbone hydrolysis are known as hemicellulases. The structure of hemicellulose reveals it as a heterogeneous polymer with different side group as a result of which large and complex enzyme groups are required for enzymatic degradation. Hemicellulolytic enzymes are classified and characterized on basis of substrate they act upon. Their modes of action on particular substrate are shown in Table 3.

Table 3. Hemicellulase enzymes and their mode of action (Jeffries 1994; Jørgensen et al. 2007).

Enzymes	**EC number**	**Mode of Action**
Exo-β-1,4-xylosidase	3.2.1.37	Release xylose from xylobiose and short chain xylooligosaccharides
Endo-β-1,4-xylanase	3.2.1.8	Hydrolyse mainly interior β-1,4-xylose linkage of the xylose backbone
Exo-β-1,4-mannosidase	3.2.1.25	Cleaves manno-oligosaccharides to mannose
Endo-β-1,4-mannasnase	3.2.1.78	Cleaves internal bonds in mannan and liberate manno-oligosaccharride
α-Galactosidase	3.2.122	Removes the galactose unit of the side chain
α-Glucuronidase	3.2.1.139	Release glucuronic acid from glucuronoxylans
Endo-galactanase	3.2.1.89	Cleaves β-1,4-galactan
Acetyl Xylan esterases	3.1.1.72	Hydrolyse acetyl ester bonds in acetyl xylans
Acetyl mannan esterase	3.1.1.6	Hydrolyse acetyl mannan bonds in acetyl mannan
Ferulic and p-cumaric acid esterase	3.2.1.73	Hydrolyse feruloyester bond and p-coumaryl ester bond in xylans
α-Arabinofuranosidase	3.2.1.55	Hydrolyse terminal nonreducing α-arabinofuranose from arabinoxylans

In nature, degradation or hydrolysis of cellulose or cellulosic biomass is executed by a set of hydrolytic enzymes jointly known as cellulases. Till date, cellulolytic enzymes are classified in three main classes: (Dashtban et al. 2009; Lynd et al. 2002).

- exo-1, 4-β-D-glucanase (EC 3.2.1.91)
- endo-1, 4-β-D-glucanase (EC 3.2.1.4)
- 1, 4-β-D-glucosidase(EC 3.2.1.21)

Cellulose can be degraded by microbes in both aerobic and anaerobic conditions. Most of cellulose in nature is degraded by aerobic system, while only 5 to 10% of cellulose in nature is degraded by anaerobic microbes liberating methane and hydrogen as end product (Carere et al. 2008). The site of action for cellulolytic enzymes including *β*-glucosidase is shown in Figure 9.

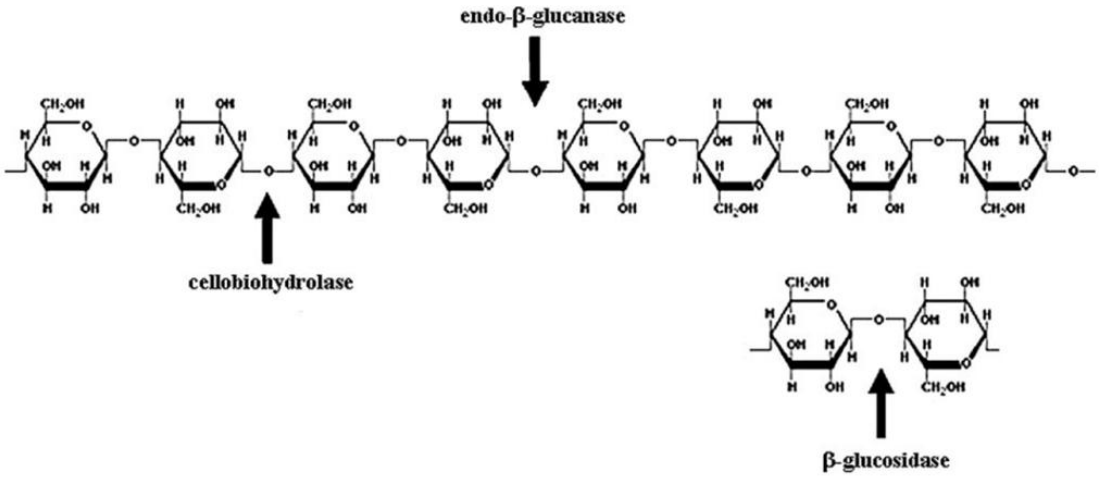


Figure 9. Site of action of three cellulase enzymes on cellulose backbone (Kumar et al. 2008)

The modular structure of cellulases reveal that they contain independently folding, structurally and functionally discrete units called domains or modules. Normally, cellulolytic enzymes consist of two domains: carbohydrate binding domain (CBD) and catalytic domain. Carbohydrate binding domain is present in C-terminal of the polypeptide connected by short poly-linker region to catalytic domain at N-terminal of the polypeptidic chain. The mode of action of cellulose hydrolysis by cellulase is either by inversion or retention of configuration of an anomeric carbon (Dashtban et al. 2009; Maki et al. 2009).

Most of the cellulases belong to the group of "glycoside hydrolases (GH) family". This family includes glycosidases and transglycosidases and is responsible for hydrolysis or transglycosylation of glycosidic bonds. More than 47% of enzymes classified in carbohydrate active enzyme database (CAZy) belong to glycoside hydrolases family

because of large variation present in genes coding glycoside hydrolases in majority of genomes. Till date, in CAZy database almost 2500 GH enzymes has been identified and sorted out into 115 families. A particular enzyme family in a CAZy data base can have different source of origin (plant, bacteria and fungi), different enzyme activity and substrate specifications (Cantarel et al. 2009; Dashtban et al. 2009). Enzymatic hydrolysis of cellulase is carried out by a set of hydrolytic enzymes which can be single enzymes (single polypeptide with multiple cellulosic domains) or extracellular multi enzyme complex. The natural occurrence of cellulase enzymes exists in two forms (Ding et al. 2008):

- Free enzyme system or non-aggregating enzymes produced mostly by aerobic bacteria and fungi.
- Aggregating enzymes systems where celluloytic enzymes form a complex often called as "cellulosome". Aggregating enzyme complexes are mostly produced in anaerobic bacteria.

3.1 Non- aggregating enzymes

In this system, cellulolytic enzymes are produced in high concentration as a single enzyme connected to binding modules and act in a synergistic manner to facilitate complete hydrolysis of cellulose *β*-1,4-glycosidic bonds to form glucose (Schwarz 2001). In general, three cellulolytic enzymes fall under this category: exo-1, 4-β-D-glucanase (EC 3.2.1.91), endo-1, 4-β-D-glucanase (EC 3.2.1.4) and 1, 4-β-D-glucosidase (EC 3.2.2.21) (Harry et al. 1993). Table 4 comprises three cellulolytic enzymes, their enzyme commission number and description for mode of action in cellulose polymer (Gowen et al. 2010).

Table 4. List of cellulolytic enzymes and their functional categories.

Enzyme category	Official name [36] (*synonym*)	Enzyme commission number	Description [36]
Endoglucanase	Cellulase (*endoglucanase*, *endo-1,4-β-glucanase*)	3.2.1.4	Endohydrolysis of 1,4-*β*-D-glucosidic linkages
Exoglucanase	Glucan 1,4-*β*-glucosidase (*exo-1,4-β-glucosidase*)	3.2.1.74	Hydrolysis of the terminal 1,4-*β*-linkages to yield one glucose
	Cellulose 1,4-*β*-cellobiosidase (*exoglucanase*, *cellobiohydrolase*)	3.2.1.91	Hydrolysis of 1,4-*β*-D-glucosidic linkages in cellulose to yield cellobiose
Cellodextrinase	*β*-Glucosidase (*cellobiase*)	3.2.1.21	Hydrolysis of terminal, non-reducing *β*-D-glucose residues to yield one glucose
	Cellobiose phosphorylase	2.4.1.20	Phosphorylates cellobiose to yield glucose and glucose 1-phosphate

3.1.1 Endo-1, 4-β-D-glucanase (EC 3.2.1.4, endocellulase)

Endoglucanases initiate hydrolysis of cellulose backbone by cleaving internal *β*-1, 4-glycosidic bond. Mostly amorphous region of cellulose is hydrolyzed by endoglucanases creating more free chain ends for cellobiohydrolase. The endoglucanase activity of cellulolytic enzymes can be assayed by using soluble cellulose substrates like carboxymethyl cellulose (Kumar et al. 2008).

3.1.2 Exo-1, 4-β-D-glucanase (EC 3.2.1.91) or cellobiohydrolase

Cellobiohydrolases are exo-acting enzyme that hydrolyzes *β*-1, 4-glycosidic bond from chain ends of cellulose backbone and produces cellobiose as main hydrolysis product. Cellobiose can act as a competitive inhibitor during cellulose degradation, resulting in retardation of cellobiose hydrolysis property of enzyme. This can result in an incomplete hydrolysis of cellobiose molecules present in the system (Dashtban et al. 2009). Exo-acting enzymes like cellobiohydrolases can form tunnel shaped closed active site for substrate binding which retains a single glucan chain and can prevent it from re adhering to the cellulose backbone (Harry et al. 1993). Figure 10 shows the difference in conformation between endo- and exo- acting cellulase enzyme.

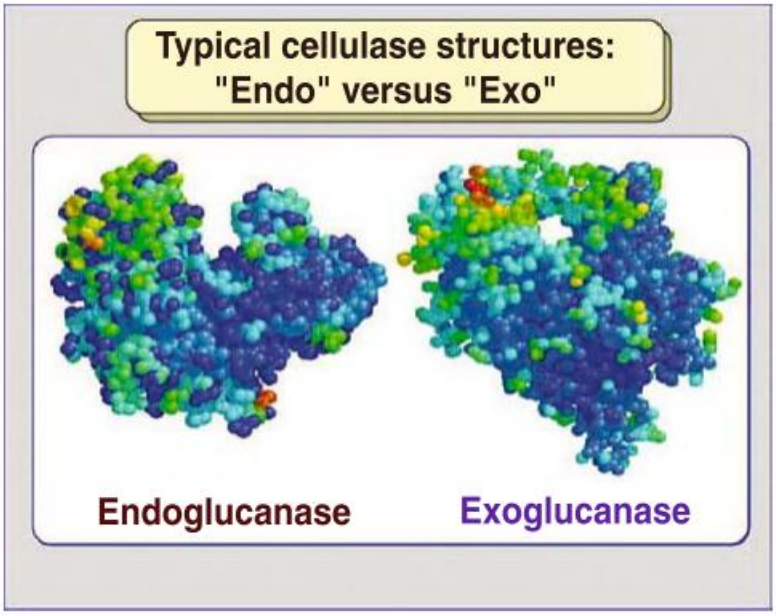


Figure 10. Conformation of endoglucanase and exoglucanase enzymes (Bayer et al. 2006).

3.1.3 β-D-glucosidase (EC 3.2.2.21)

The third essential enzyme for cellulose hydrolysis is *β*-D-glucosidase. This enzyme hydrolyzes β-1, 4 linkages from soluble cellobiose and cello-oligosaccharides and releases two glucose monomers. In some cases this enzyme can also act on large substrate molecules like cellotriose (Gowen et al. 2010).

Glucose molecules act as a competitive inhibitor of *β*-glucosidase (Dashtban et al. 2009). In CAZy database, *β*-D-glucosidases have been put in family 1 and 3 of glycoside hydrolases based on amino acid sequence. All *β*-D-glucosidases from various source of origin like bacteria, plant and fungi are kept in family 3, whereas family 1 glycoside hydrolases contains *β*-glucosidase from mammalian, plant and bacteria origins which posses dual enzymatic activity: galactosidase and *β*-glucosidase activity (Dashtban et al. 2009).

3.2 Aggregating systems

In aggregating enzyme system, cellulases produced from celluloytic microorganism associate to form a multi-enzymatic complex called cellulosome (Ding et al. 2008). Cellulosome allows concentrated enzyme activity in close proximity of the bacterial cells. Moreover, aggregated enzymes show a higher hydrolytic efficiency compared to non-aggregated cellulase (Gowen et al. 2010; Kumar et al. 2008; Lynd et al. 2002). The cellulosome in cellulolytic microorganism are found to be associated with cell surface. It facilitates the attachment between cells and insoluble substrate along with efficient uptake of hydrolysis products by host cell and also prevents loss of hydrolysis product by diffusion or uptake by other microbes (Gowen et al. 2010; Harry et al. 1993).There exists inter- and intra- species divergence between cellulosome's composition. Interspecies variation in cellulosome depends upon the properties of scaffoldin protein, whereas intraspecies variation is due to the type of enzyme that binds to the scaffoldin protein. Some bacterial cellulosome contains single type of scaffoldin protein, whereas others have multiple scaffoldin which results in variation of cellulosome (Doi et al. 2004).

In general, all cellulosomes contain some common protein subunits. These subunits include fibrillar proteins (scaffoldin proteins) that contain cohesion binding site for interacting with celluosomal enzyme subunits. Cellulosomal subunits also contain different functional and invariable regions called dockerin for binding and interaction with cohesin of scaffoldin protein. The most important factor for cellulosome assembly is the cohesin and dockerin interaction. All dockerin domains of cellulosomal complex interact with cohesin of scaffoldin proteins (Doi et al. 2004). The scaffoldin proteins play major role in cellulosome as they carry out three major functions like binding to cellulosomal enzymes,

binding to cellulose substrate and binding cell surface associated proteins. Apart from cohesin active site, the scaffoldin also contains carbohydrate binding modules (CBM) or carbohydrate binding domains (CBD) that bind to cellulose and hold it during enzymatic hydrolysis (Ding et al. 2008; Gowen et al. 2010; Shoham et al. 1999). Figure 11 represents the structure and protein subunits of typical *C. thermocellum* cellulosome.

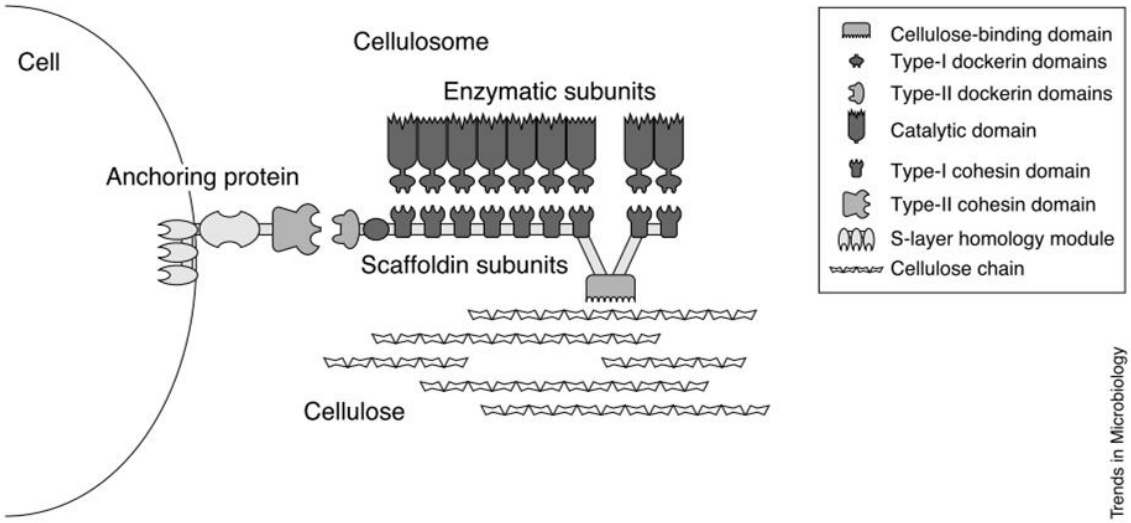


Figure 11. *Clostridium. thermocellum* cellulosome and its protein subunits (Shoham et al. 1999).

Some of the major advantages of aggregating cellulolytic enzyme systems (cellulosome) over non-aggregating cellulolytic enzymes for efficient hydrolysis of cellulose are listed below (Maki et al. 2009; Schwarz 2001).

- Eliminates expenditure of energy for producing copious amount of free enzymes, which has possibility to get diluted and lost in bulk solution.
- Helps in optimization of synergism by correct ratio between the components, which is determined by composition of the complex.
- Helps in elimination of non productive adsorption by optimal spacing of components working together in synergistic style.
- It facilitates in binding whole enzyme complex to single site of biomass surface through strong domain with low specificity and eliminates competitiveness in binding for active site of biomass.
- Close proximity of enzyme substrate interaction decreases intermediate transit time and increases catalytic efficiency. Performs close monitoring of inhibitory products and mediates passage of cellobiose and cellodextrins inside cell for metabolism.

- Finally, depletion of one structural cellulose type at site of adsorption can halt or interrupt overall hydrolysis process. In cellulosome, this kind of pause can be avoided by occurrence of several other enzymes of different specificity.

3.3 Sources of cellulolytic enzymes

Cellulases has been reported from a diverse range of cellulolytic microorganism living in extremely diverse environmental conditions which includes soil samples, fresh and salt water sediments, thermal springs, extremely cold environment and also as symbionts in terrestrial and marine organisms. These microorganisms mainly include fungi, bacteria and yeast which have been studied for several years (Gowen et al. 2010; Howard et al. 2003; Lee 1997). As now, no single microorganism studied has been reported capable of producing cellulolytic enzymes for efficient biodegradation or hydrolysis of cellulose biomass. Efficient biodegradation of cellulosic biomass in nature is carried out by microorganisms in cooperation. This microbial cooperation can occur between different fungal species and bacterial genera that produce diverse celluloytic and hemicellulolytic enzymes in different aerobic and anaerobic growth conditions. The mutualism between different microbes can avoid various problems like feedback regulation and metabolite repression that can occur in single strain (Wongwilaiwalin et al. 2010). Haruta et al. 2002 have reported that symbiosis of one celluloytic bacterium with other non-celluloytic bacteria can be an ideal condition for efficient cellulose degradation.

Secretion of cellulolytic and hemicellulolytic enzymes has been so far reported from both prokaryotic and eukaryotic microorganisms. In eukaryotes, fungi have been studied extensively. Especially, aerobic fungal strains like *Trichoderma reesei* and *Aspergillus Niger* posses the capacity to secrete high concentration and variety of cellulolytic enzymes. *T. reesei* was the first cellulolytic microbes isolated in 1950's. Extensive studies and strain improvement have developed efficient mutants that are used in various industrial and biofuel applications (Howard et al. 2003; Shallom et al. 2003; Tengerdy et al. 2003).

Apart from fungi, several bacterial species have been studied for cellulase production. Both aerobic and anaerobic bacteria have the ability to produce cellulolytic enzymes. In comparison to *T. reesei* aerobic bacteria like *Bacillus* sp. and *Cellvibrio sp.* secrets only a small pool of cellulose and hemicellulose degrading enzymes. However, they are still considered as strain of interest due to their unique lignocellulolytic gene pools (Shallom et al. 2003).

Like aerobes, anaerobic bacteria also secrete cellulolytic enzymes and have received a lot more attention because of their complex enzyme system. Anaerobic bacteria synthesize cellulosomes which can incorporate different cellulases and hemicellulases for the hydrolysis of biomass

3.4 Thermophilic bacteria

Thermophilic microorganisms can be found in phyla of bacteria and archaea. The current nomenclatures of thermophilic microbes are based on optimum growth temperature and can be called as moderate thermophiles (50-70 °C), extreme thermophiles (≥70 °C) and hyperthermophiles having optimum growth temperature at 80 °C or above. Mostly, extreme thermophiles and hyperthermophiles grows in an extreme environment like terrestrial hot springs, solfataric fields, shallow submarine hydrothermal system, geothermally heated oil reservoirs. Most of thermophilic bacteria are found to be anaerobic in nature due to low oxygen solubility at high temperature and also presence of reducing gases in extreme environments (Marchant et al. 2002; VanFossen et al. 2008).

Thermophilic bacteria have developed a unique system to withstand high temperature and to carry out their cellular activity. Their major adaptation in membrane lipid to withstand extreme temperature includes modification of the phospholipids bilayer by increasing acyl chain length, increase in degree of saturation of fatty acid and insertion of branching and cyclization in lipid layer (Gomes et al. 2004). The structural and functional integrity of the nucleic acid to withstand an extreme temperature is maintained by increasing negative or positive supercoiling in nucleic acid, with the help of reverse DNA gyrase enzyme that introduces positive supercoils in the nucleic acids (Charlier et al. 2005). Thermophilic microbes produce special protein molecules known as chaperonins which helps in protein refolding to their native state and restoration of protein function for cellular activities. As a result of which proteins from thermophilic microbe are resistant to denaturation and proteolysis (Haki et al. 2003) .

Interest in study of thermophilic microbes over past 15-20 years has increased dramatically because of the importance of thermostable enzymes (e.g. DNA polymerase), their fermentation products and their function in nature. Most of research work on thermophiles has been directed towards biotechnological applicable microbes which have potential importance in fermentation process, cellulose and hemicellulose degradation and biofuel generation (Bredholt et al. 1995). Use of thermophilic bacteria to carry out fermentation process at high temperature provides additional advantages. The high temperature fermentation process can reduce production cost; reduce risk of contamination by common mesophiles, and also product inhibition. Also, thermophilic microbes are noted

to have a metabolic diversity in degrading and utilizing various unprocessed carbohydrate source in comparison to various industrial mesophilic organisms like yeast (Weber et al. 2010). Table 5 includes list of some anaerobic thermophilic bacteria, their genome size and growth optima.

Table 5. Anaerobic extreme thermophiles with sequenced genome (VanFossen et al. 2008).

Microorganism	Domain	Size (Mb)	Growth T (°C)	Metabolism
Caldicellulosiruptor saccharolyticus	Bacteria	2.97	70	Heterotroph
Thermoanaerobacter tengcongensis	Bacteria	2.68	75	Heterotroph
Thermotoga maritima MSB8	Bacteria	1.86	80	Heterotroph
Thermotoga neapolitana	Bacteria	~1.8	80	Heterotroph
Thermotoga petrophila	Bacteria	1.82	80	Heterotroph
Archaeoglobus fulgidus	Archaea	2.17	83	Heterotroph / Autotroph
Methanococcus jannaschii	Archaea	1.73	85	Autotroph
Pyrococcus horikoshii	Archaea	1.73	95	Heterotroph
Hyperthermus butylicus	Archaea	1.67	95–106	Heterotroph
Methanopyrus kandleri	Archaea	1.69	98	Autotroph
Staphylothermus marinus	Archaea	1.57	98	Heterotroph
Pyrococcus abyssi	Archaea	1.76	100	Heterotroph
Pyrococcus furiosus	Archaea	1.90	100	Heterotroph
Pyrobaculum islandicum	Archaea	1.83	103	Heterotroph

3.5 Thermostable enzymes and cellulase

The industrial application of thermostable enzymes is increasing day by day due their resistance to harsh conditions typical of industrial processes. Most of cellular components like nucleic acid, proteins and enzymes from thermophilic organisms are known to be stable at high temperature, resistant to denaturants and active at extreme acidic or alkaline conditions (Haki et al. 2003).

Thermostable enzymes can remain stable and active even above the optimum growth temperature of microorganisms (Haki et al. 2003). However, exact concept of thermostability is still unclear. An enzyme or protein can be defined as thermostable only if posses high defined unfolding (transition) temperature (T_m) or possess a long half life of enzymatic activity at elevated temperature. In most cases, enzyme activity increases with an increase in temperature till it reaches inactivation temperature (Turner et al. 2007; Viikari et al. 2007).

Thermostable enzymes can be produced from both mesophilic and thermophilic microbes; however, most of thermophilic enzymes used in industrial process originate from mesophilic microbes (Viikari et al. 2007). The current demands of thermostable enzymes are in food, textile, starch, leather, pulp and paper and pharmaceuticals industries. Among them starch industry is one of the largest consumer of thermostable amylases like α-amylases, β-amylase, glucoamylases and iso-amylases or pullulanases that are used for hydrolysis and modification of raw starch to sugars and various other products (Gomes et al. 2004). The group of starch hydrolytic enzymes accounts for 30% of world's total enzyme consumption (Haki et al. 2003).

On the other hand, most of cellulolytic enzymes are used in industries for color extraction of juices, in detergents, bistoning of jeans, and pretreatment of biomass that contains cellulose. Many thermostable cellulases characterized till date has been experimented in these areas because of the high performance of thermostable enzymes (Turner et al. 2007). In biorefining area, cellulose biomass is converted to fermentable sugars for biofuel generation with the help of cellulolytic enzymes. However, low hydrolytic activity and expensive production cost for large scale application of existing cellulases has directed an interest for economic cellulase production and with better enzymatic activity on cellulose biomass. This resulted in search for novel cellulase from thermophilic bacteria and has received great attention for highly active and thermostable cellulases (Viikari et al. 2007; Wang et al. 2010).

Studying and characterization of thermophilic cellulase enzymes from thermophiles is increasing mainly because of possibility to clone and express it in *Bacillus subtilis* and *E. coli* (Zamost et al. 1991). *C. thermocellum* belonging to bacterial genus *Clostridium* is one of most broadly studied rod shaped and spore forming cellulolytic thermophiles (VanFossen et al. 2008). Currently, several other cellulase producing anaerobic thermophilic bacteria strains like *Caldicellulosiruptor* sp. and *Thermotoga* sp. have driven an attention in search of thermostable free acting cellulase that are not part of cellulosomal complex. Hydrolysis of cellulose in lignocellulose biomass by thermostable cellulase provides several benefits (Viikari et al. 2007).

- Thermostable cellulase has higher specific activity as a result of which less enzyme is required for hydrolysis reaction.
- Highly stable thus can carry out biomass hydrolysis for extension time then conventional enzymes.
- Improved performance and decreases overall hydrolysis cost.

3.6 *Caldicellulosiruptor bescii*

Lignocellulosic biomass deconstruction for biofuel production by thermophilic microbes (Topt ≥70 °C) has renewed interest over extreme thermophiles that grow around temperature range of $T_{opt} \geq 80$ °C. Till date, hyperthermophilic bacteria capable of lignocellulose hydrolysis have not been reported. On the other hand, bacterial strain having the optimal growth temperature in the range of 70 °C to 80 °C for plant biomass degradation has been isolated and identified (Blumer-Schuette et al. 2010).

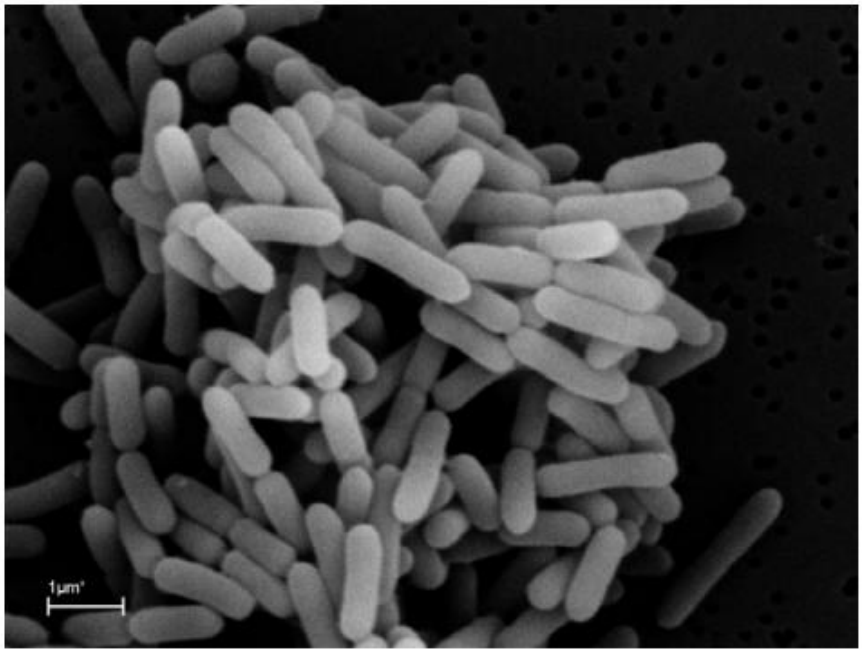


Figure 12. *Caldicellulosiruptor bescii* cell morphology under scanning electron microscope (Yang et al. 2010b).

C. bescii is a gram positive and asporogenic bacteria having optimal growth temperature of 75 °C. *C. bescii* formerly named *Anaerocellum thermophilum* was initially isolated from a Russian hot spring and deposited in DSMZ culture collection as DSMZ 6725 based on its phenotypic and physiological characteristics in absence of 16s RNA gene sequence. Reclassification and 16S rRNA sequence study of strain DSMZ 6725 revealed that it falls under *Caldicellulosiruptor* clade and hence *A. thermophilum* has been reclassified as *C. bescii*. Recently, genome sequencing of *C. bescii* has been completed and now available (Yang et al. 2010b).

C. bescii has a potential role in development of second generation biofuel production from lignocellulose biomass because of its high optimum growth temperature and secretion of cellulose and hemicellulose degrading enzymes into extracellular medium which are not part of cellulosome complex found commonly in *Clostridia* (Blumer-Schuette et al. 2010).

3.7 Screening and assay methods for cellulase

In nature, various microbes including fungi and bacteria are known to produce cellulolytic enzymes for cellulose hydrolysis. However, complexity arises in finding particular method that can be used to detect cellulase producing microbes and quantifying produced cellulase activity. The most widely used technique for screening cellulase producing bacterial species follows plate screening methods. In the plate screening methods, several dyes (Congo red, Ruthenium Red, Calcofluor White, and Iodine) interacting with polysaccharide (cellulose) specifically or non-specifically are used. In addition, cellulase producers can be detected upon polysaccharide degradation in culture medium during growth. Later the enzymatic degradation and consumption of the polysaccharide can be also visualized as clear halo zone in gel matrix (Badel et al. 2011; Dashtban et al. 2010).

3.7.1 Plate screening method

The most widely used technique to detect celluloytic bacteria in an agar plate is carried out using Congo red dye where microbes are plated in gel matrix containing cellulose as a substrate. The plate with gel matrix are flooded with Congo red dye and incubated for interaction to occur. The unbound dye from plate is washed off and can be visualized to find cellulose degrading microbes. The area in plate containing unstained spots confirms the polysaccharide (cellulose) degradation by bacteria colony due to celluloytic activity (Ruijssenaars et al. 2001).

The main drawback of Congo red technique is that it can detect only those microbes that depolymerize polysaccharide by endocellulase (that cleaves cellulose backbone randomly), as a result of which microbes producing exo-acting (that removes only few glucose molecules from the chain end) cellulase enzyme for polysaccharide hydrolysis can escape detection by this technique. Apart from plate screening method, Congo red has also been used to detect celluloytic enzyme activity in gel electrophoresis (zymogram technique) and also in gel diffusion assay of enzyme fraction (Ruijssenaars et al. 2001).

3.7.2 Celluloytic enzyme assay

Hydrolysis or degradation of cellulose polymer by cellulase is known as cellulolytic activity. Celluloytic activity can be measured through two approaches (Zhang et al. 2006)

- Measuring individual cellulase (endoglucanase, exoglucanase, and β-glucosidase) activity.
- Determining saccharifying activity of a crude cellulase system.

The major enzymes of cellulolytic systems are endoglucanases. They hydrolyze intramolecular β-1, 4-glucosidic linkage of cellulose backbone. Various substrates can be used for endoglucanase assay depending upon solubility in water, like cotton linter or Whatman No. 1 filter paper as insoluble substrate and soluble substrates like carboxymethyl cellulose (CMC). Most of the quantitative assays for cellulase activity are done on following basis (Dashtban et al. 2010):

- Product accumulation after hydrolysis.
- Reduction in substrate quantity.
- Alteration in physical property of substrates.

Quantitative assay for endoglucanases and other cellulolytic enzymes are done on the basis of hydrolysis products formed, which includes reducing sugars, total sugars and chromophores produced from substrate molecules. Insoluble substrate like Whatman No 1 filter paper and soluble substrate like CMC are widely used for quantitative assay of endoglucanases and reducing sugar produced are determined with the help of DNS reagent method (Ghose TK 1987; Miller 1959). Phenol-sulphuric acid or anthrone- sulphuric methods are used for measurement of total sugar produced during the cellulase assay. Sugar detection range in most of cellulolytic enzyme assays can be modified either diluting the color reaction solution or changing sugar volume per sample prior to reaction (Zhang et al. 2006).

4. Materials and Methods

4.1 Bacterial strains, Plasmid and growth conditions

The gram positive bacteria *C. bescii* strain DSM 6725, obtained from DSMZ (German collection of Microorganisms and Cell Cultures) was grown in DSMZ medium 516 (http://www.dsmz.de/microorganisms/media_list.php). Preparation of growth medium was carried out by following the instructions of DSMZ medium 516. The medium was prepared anaerobically under a N_2-CO_2 (80:20) atmosphere and sterilized separately. Later, anaerobic stock solution of bicarbonate, cellobiose and sodium sulfide were added. Final pH of the medium was adjusted to 7.2. Cells were grown as static culture in 50 ml falcon tubes kept inside an anaerobic jar incubated at 75 ℃. Bacterial cells were incubated for 3 days until noticeable growth. Genomic DNA was extracted using E.Z.N.A Bacterial DNA isolation kit (Omega Bio-tek, USA) by following manufacturer's protocol. DNA was quantified by measuring the absorbance spectrophotometrically at 260 nm (Gene Quant pro, Amersham Biosciences) and stored at – 20 ℃.

E.coli cells carrying plasmid pVKK81 (3030 bp) was inoculated from glycerol stock to 5 ml LB medium containing 0.5 % glucose and 12.5 µg/ml of tetracycline antibiotic. The culture tubes were incubated at 37 ℃ /250 rpm for cell growth. After overnight incubation, 15 µl of the pre-culture was inoculated to 250 ml conical flask containing 30 ml LB, 12.5 µg/ml of tetracycline and 0.5 % glucose. The cells were grown at 37 ℃ /250 rpm. Once optical density (OD) of cells at 600 nm reached 0.6, plasmid extraction was carried out using GenElute plasmid mini prep kit (Sigma-Aldrich, USA). The purity and plasmid DNA concentration was measured spectrophotometrically as previously mentioned. .

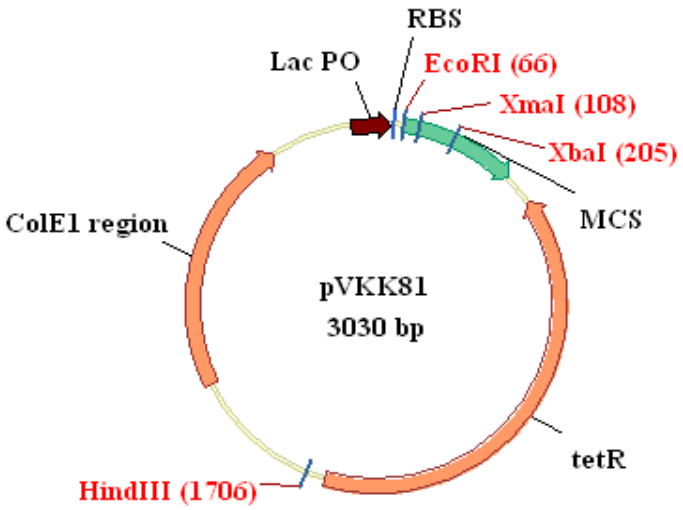


Figure 13. Map of cloning vector pVKK81 (3030 bp).

Table 6. Bacterial strains, plasmid and gene used in this study.

Strains and plasmid	**Description**	**Source**
E.coli BL21	Expression host strain	From our laboratory.
C. bescii	Thermophlic known as cellulose degrader	DSMZ 6725.
Plasmids		
pVKK81	General cloning and expression vector, Tet^{res}	Obtained from our lab.
pSB01-endo	Endoglucanase gene cloned into pVKK81	Constructed in this work.
pSB02-exo	Exoglucanase gene cloned into pVKK81	Constructed in this work.

4.2.1 Nucleotide sequence and accession number

Nucleotide and amino acid sequence of endoglucanase from *C. bescii* are available in the GeneBank with Gene ID of 7406935 and for cellobiohydrolase with Gene ID of 7407174. The genome of *C.bescii* contains a circular chromosome of 2919718 bp with 35.2% GC content and 2666 protein coding sequences organized into 1209 operons. The average length of protein coding gene in the chromosome of *C .bescii* is found to be 942 bp. Altogether, 394 (14.8%) proteins out of 2666 proteins encoded in the chromosome of *C. bescii* are predicted to have signal peptide and 344 (12.9%) proteins predicted to have transmembrane helices (Dam et al. 2011).

Table 7. General features of *C. bescii* DSM 6725 (Dam et al. 2011).

General features	***C. bescii* (DSM 6725)**
Length of Chromosome (Mbp)	2.9
G+C content (%)	35.2
Coding density (%)	85.4
Total no. of predicted protein coding genes (bp)	2662
Number of Secreted proteome (SignalP prediction)	394
Average length of protein coding genes (bp)	942
Growth on cellulose and xylan	Cellulose, xylan

4.3 Gene Construction

4.3.1 Construction of Endo-l, 4-β-D-glucanase expression vector

The DNA fragment encoding endo-l, 4-β-D-glucanase gene was amplified from genomic DNA of *C. bescii* as template with primers Sb-01 and Sb-02 (in detail Table 10). Ribosomal binding site (RBS) was inserted in forward primer after *Eco*RI site. These primers amplified the whole sequence of endo-l, 4-β-D-glucanase (Figure 15) including the signal peptide. The amplified PCR product was restricted and ligated to the cloning vector (pVKK81) as shown in Figure 15. The cloned gene in plasmid (pVKK81) was under regulation of inductive lac promoter (T_4LacPO) and having tetracycline resistance gene ($Tet^{res)}$ as a selection marker. PCR, restriction and ligation methods are discussed briefly under the section 4.4.

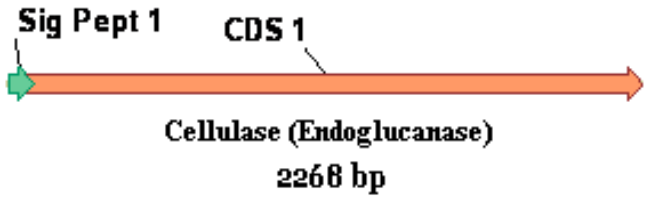


Figure 14. Map of cellulase (endoglucanase) gene from *C. bescii*.

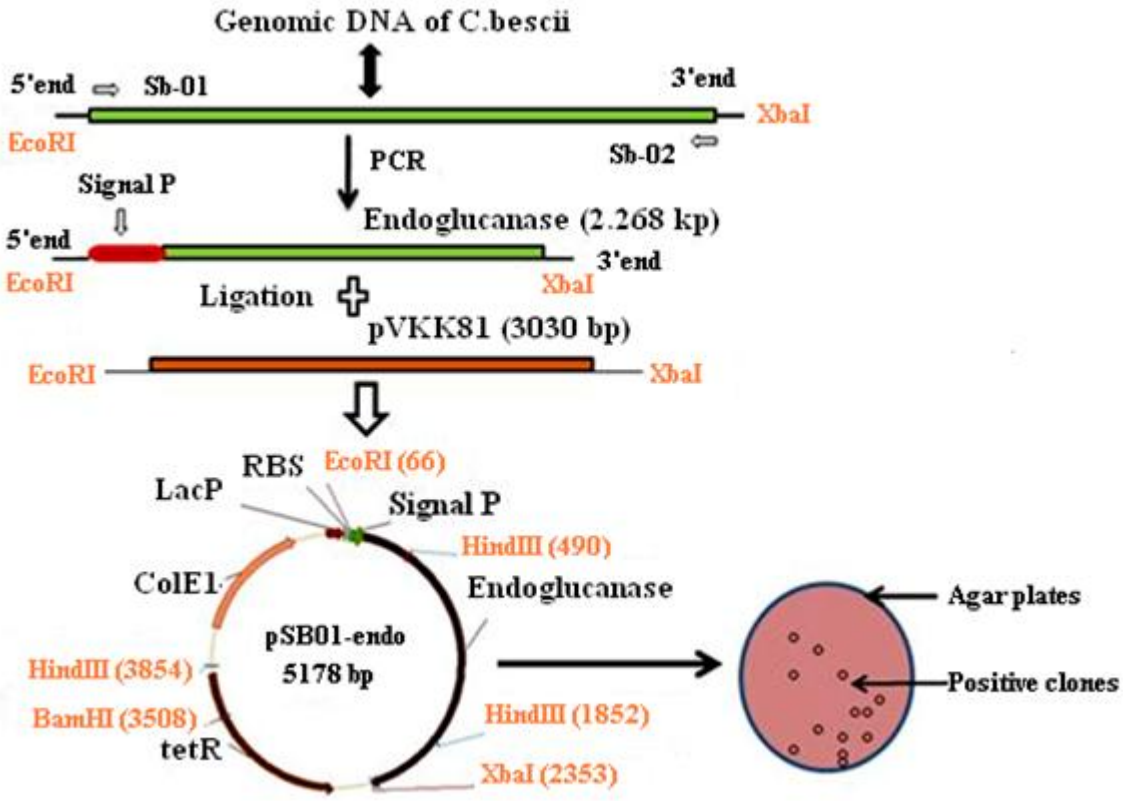


Figure 15. Scheme depicting the construction of plasmid pSB01-endo

4.3.2 Construction of exoglucanase expression vector

Figure 16 shows the strategy employed for cloning exoglucanase gene in plasmid pVKK81. PCR amplification of exoglucanase gene including the signal peptide was carried out using genomic DNA of *C. bescii* as template with primers Sb-03 and Sb-04 (Table 10). Ribosomal binding site (RBS) was inserted in the forward primer after *Eco*RI site. The cloned exoglucanase gene is also under the regulation of inductive lac promoter (T_4LacPO) in pVKK81 and having tetracycline resistance gene (Tet^{res}) as a selection marker.

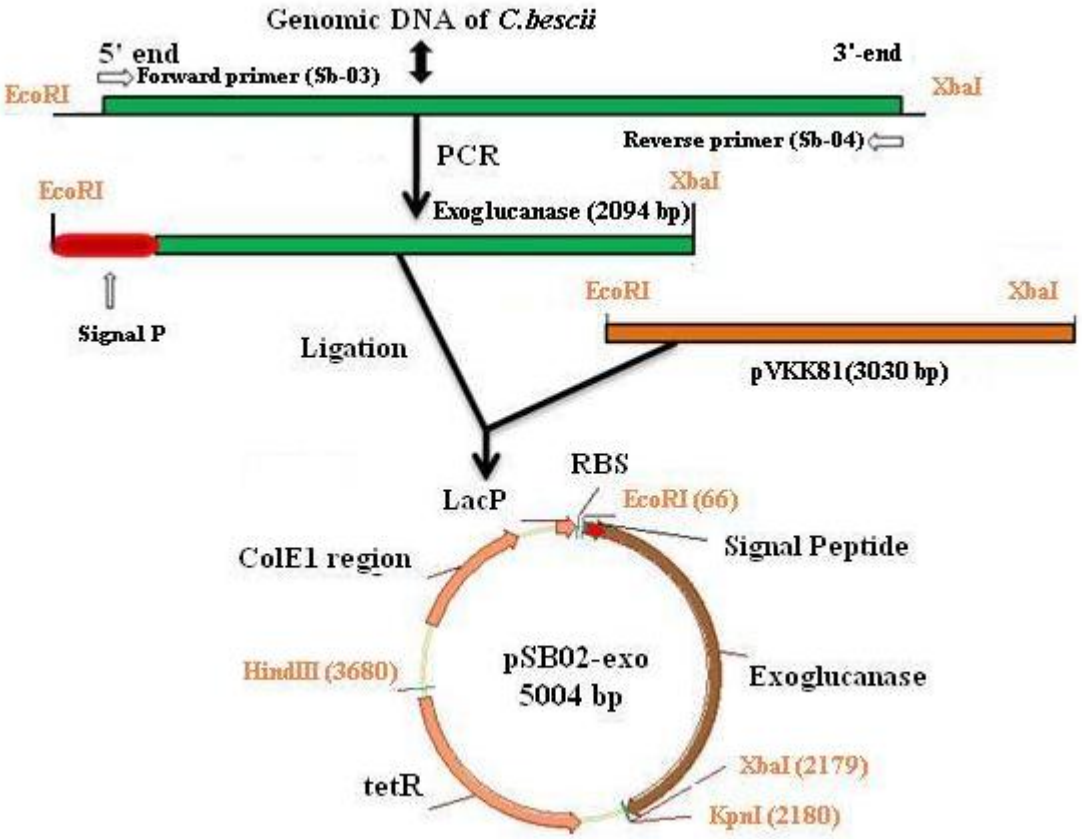


Figure 16. Flow chart for construction of plasmid pSB02-exo.

4.4 Gene Manipulation

4.4.1 PCR Amplification

Amplification of cellobiohydrolase and endoglucanase genes from genomic DNA of *C. bescii* was carried out by polymerase chain reaction (Image in Appendix 1 and 2). Following reaction was routinely used to set up PCR reaction: 100-200 ng of genomic DNA, 0.25 µmol L^{-1} sense and antisense primers (Table 10), 1x Optimized DyNAzyme™ Buffer, 0.5 U DyNAzyme™ II PCR polymerase, 0.2 mm mol L^{-1} deoxynucleotide triphosphate (dNTPs) and sterile- Dnase free Milli-Q water to make up the reaction mixture to 50 µl. PCR reaction was carried out in T3000 Thermocycler using PCR programs mentioned in Table 8 and 9 for respective genes. After completion of the PCR reaction, success of target gene amplification was confirmed by performing gel electrophoresis of samples in 1% (w/v) agarose gel.

Table 8. PCR program used for amplification of exoglucanase (cellobiohydrolase) gene.

Program and temperature	**Time**
1. Initial denaturation 94 °C	4 minutes
2. Denaturation 94 °C	30 seconds
3. Annealing 50 °C / 55 °C	1 minutes
4. Elongation 72 °C	2 minutes
5. Final elongation 72 °C	10 minutes
6. Pause 4 °C	
Number of Cycles - 30	

Table 9. PCR program used for the amplification of endoglucanase gene (2268 bp).

Program and temperature	**Time**
1. Initial denaturation 94 °C	4 minutes
2. Denaturation 94 °C	30 seconds
3. Annealing 57 °C / 60 °C	1 minutes
4. Elongation 72 °C	2 minutes
5. Final elongation 72 °C	10 minutes
6. Pause 4 °C	
Number of cycles – 30	

Table 10. Primer sequence and description. Bold alphabets represent the endonucleases site, and ribosomal binding sites whereas start and stops codons are represented by lowercase bold alphabets.

Primer	**Description**	**Oligo sequence (5'→3')**
Sb-03	Cellobiohydrolase forward	ATTA**GAATTC**GA**AGGA**GATATCC**ATG**AAAAAAAGGAAA TTCAAAATATTATATTTA
Sb-04	Cellobiohydrolase reverse	ATTA**TCTAGAtta**TTTTTTAGCCTTTACTTTTGGAATAGC
Sb-01	Endoglucanase forward	ATTA**GAATTC**GA**AGGA**GATATCC**ATG**AGGAAAATTATTT TAAAGTTTTGTGCA
Sb-02	Endoglucanase reverse	ATTA**TCTAGAtca**TAGCTTGCCTGCTAAGTTCAA**GAG**

4.4.2 Digestion and Ligation reaction

Target genes (endoglucanase and exoglucanase) and plasmid (pVKK81) were double digested with restriction endonucleases (*Xba*I and *Eco*RI) for 2 hrs at 37 °C. Digested fragments were purified using 1 % (w/v) agarose gel electrophoresis and correct DNA fragments were excised out and purified using Gene JET^{TM} Gel Extraction Kit (Fermantas, Finland) by following manufacturer's protocol. Purified linear vectors and target genes were ligated at 16 °C overnight using T4 DNA ligase.

4.4.3 Transformation in *E. coli* (BL21)

Electrocompetent *E.coli* (BL21) cells were prepared using standard protocols (Sambrook et al., 1990). Ligation mixture of 2 µl and 50 µl of competent cells were mixed together and transferred to chilled electroporation cuvette. Transformation of cells was carried out in Gene Pulser MXcell electroporation system (Bio-Rad, USA) using Eco-1 program. After electric pulse, pre warmed LB media (1 ml) was added to electroporation cuvette and cells were incubated at 37 °C with gentle shaking. Transformed cells were spread on pre-warmed LA plates containing 12.5 µg/ml of tetracycline antibiotic. The plates were incubated overnight at 37 °C.

4.5 Screening methods

Screening and conformation of positive clones carrying endoglucanase and exoglucanase genes after transformation in *E.coli* BL21 was carried out by PCR method, restriction analysis and also by Congo red assay.

4.5.1 PCR and restriction analysis

PCR was used to confirm insert gene inside cloning vector using the plasmid DNA extracted from the transformants as template. Agarose gel electrophoresis was used to analyze the amplified PCR products of correct base pair. PCR program and reaction conditions are mentioned in section 4.4.1. Restriction analysis of extracted plasmid DNA was also carried out to confirm cloning of target genes in the vector. After restriction analysis, the presence of target genes was confirmed by agarose gel electrophoresis (Appendix 3 and 4).

4.5.2 Congo red assay

The CMC/Congo red method (Ruijssenaars et al. 2001) was utilized to screen the positive clones harboring endoglucanase and exoglucanase genes. The positive clones obtained from overnight incubation after transformation were streaked out in LA agar plates containing CMC 0.5 %, tryptone (1%), yeast extract (0.5 %) , sodium chloride (1%) and 1.5% agar. The plates were incubated overnight at 37 ℃ for cell growth. After incubation, the plates were flooded with an aqueous Congo red solution (1g/l) (Sigma-Aldrich, USA) and incubated in room temperature for 10-15 minutes. Subsequently, the Congo red solution was poured off and replaced with 1 M sodium chloride. The plates were again incubated for 10-15 min in room temperature to maximize the binding of dye to polysaccharide CMC. Finally, the plates were washed with several round of distilled water to remove all unbound dye. The positive clones harboring the cellulolytic genes showed non or lightly colored haloes on CMC plates and this indicated the degradation of CMC polysaccharide. Finally the image were taken and labeled.

Apart from Congo red assay, cellulase activity from the screened clones was also confirmed by the DNS test method using crude protein extracts and culture supernatant of the cell culture. Procedure for DNS assay is mentioned in section 4.7.3.

4.6 Enzyme (protein) expression and purification

4.6.1 Inoculum preparation

The *E.coli* BL21 cells harboring plasmid pSB01-endo were grown in 30 ml of 2XYT medium (tryptone 1.6%, yeast extract 1%, NaCl 0.5%) in 100 ml sterile conical flask containing tetracycline (10 µg/ml) as a resistance marker. The cells were grown overnight at 30 ℃ / 300 rpm. *E.coli* BL21 cell as control was also grown exactly in same condition without using antibiotics in the medium.

4.6.2 Expression of endoglucanase gene in *E.coli* BL21

E.coli BL21 harboring plasmid pSB01-endo was cultured in 200 ml of sterile 2XYT medium (10µg/ml of tetracycline), and inoculated with overnight cultures keeping initial OD 0.001. *E.coli* BL21 without the plasmid was also cultured under same condition and this acts as a control. Cells were grown at 37 ℃ / 300 rpm and OD of culture was monitored in regular intervals by taking the absorbance at 600nm. Once the cultures reached early log phase (OD_{600}= 0.6), expression was induced by addition of 0.4 mM isopropyl β-D-thiogalactosidase (IPTG) for 4 hrs in same growth condition. After 4 hours, the temperature was lowered to 25 ℃ and the culture was grown overnight at 150 rpm.

4.6.3 Isolation and partial purification of crude enzyme

The overnight induced cells were harvested by centrifugation at 10,000 rpm/ 20 minutes at 4 ℃ and resuspended in 20 ml of 50 mM phosphate buffer (pH 6.0), whereas the supernatant of cell culture after cell separation was stored in sterile bottles at 4 ℃ for enzyme activity test. The resuspended cells were treated with lysozyme (1mg/ml) for 30 min at 4 ℃ prior to cell disruption by sonication on ice (2 times for 30 second each on ice). The cell lysate was incubated at 60 ℃ for 30 min and centrifuged at 10,000 rpm for 35 minutes at room temperature to remove all heat labile proteins present in cell mixture. Endoglucanase protein fraction in the supernatant of cell lysate was transferred to sterile tubes and used as a crude enzyme source for enzyme assay and enzyme characterization.

4.6.4 Molecular mass determination by SDS-PAGE

The crude enzyme extracted from supernatant of cell lysate was used as protein sample for SDS-PAGE assay. The protein samples were denatured in sample buffer containing 2.5 % glycerol, 0.5 M Tris-HCl (pH 6.8), 10 % (w/v) SDS, 0.5% (w/v) bromophenol blue and 5% β-Mercaptoethanol at 95 degree for 5 min. Different sample volume 30 µl and 20 µl was loaded on 10% resolving gel of SDS-PAGE. The gel was run at 160 voltages for about 60 minutes and protein bands were visualized by Commas brilliant blue R-250 staining.

4.7 Endo-1, 4-β-Glucanase activity assay

Endoglucanase (*EC.3.2.1.4*) hydrolyzes CMC (high viscosity) to produce free glucose units during the chemical reaction. These free glucose units react with 3, 5 Dinitrosalicyclic acid to form a colored complex which can be detected by spectrophotometer at 540 nm.

4.7.1 Carboxymethyl Cellulose (CMC) / substrate solution

One percent carboxymethyl cellulose was used as substrate solution in entire enzymatic assay prepared in different buffers. For pH and temperature experiment it was prepared in different buffers of 50 mM concentration each. Phosphate buffer, Na-Acetate buffer and Glycine buffer was used in pH assay, whereas for temperature assay only phosphate buffer was used.

4.7.2 Preparation of Ditrosalicyclic Acid (DNS) reagent

Ditrosalicyclic acid was used as reagent for determination of reducing sugars liberated by endoglucanase enzyme from carboxymethyl cellulose substrate during enzymatic reaction. DNS reagent was prepared by adding 3, 5 Dinitrosalicyclic acid (10.6 g/L), Sodium hydroxide (19.8 g/L) and Rochelle salts (Na-K tartarate) (306 g/L) in sterile MQ water and stored in dark bottle to protect from light.

4.7.3 Ditrosalicyclic acid assay for determination of enzyme activity

Determination of enzyme activity of endoglucanase and control protein was carried out using appropriate amount of endoglucanase enzyme (125 µl) and control protein (125 µl) in presence of 0.5 ml of 1% CMC prepared in buffer solutions and the final volume was adjusted to 1 ml by using appropriate buffer. The sample mixture was incubated at different temperatures for 60 min depending upon the experiment. The enzyme activity of reaction mixture was terminated by adding 3 ml of DNS reagent. The tubes were kept in boiling water bath for exactly 5 minutes for color development, followed by cooling on ice. The release of reducing sugar from substrate during enzymatic reaction was determined by diluting the reaction mixture with 20 ml of MQ water. From this mixture, 1 ml of sample solution was taken in plastic cuvette and absorbance was measured spectrophotometrically at 540 nm. The absorbance value obtained were later converted to glucose concentration and finally to enzyme units as explained in detail in section 4.8.

4.7.4 Optimum temperature, pH and thermostability determination

The optimum temperature of endoglucanase enzyme was determined by running dinitrosalicyclic assay in temperatures range from 40 °C to 90 °C for sample incubation. The assay was carried out in 25 ml glass tubes with 1 ml reaction mixture containing 0.5 ml of 1% CMC prepared in 50 mM phosphate buffer and 250 µl of crude enzyme (2 fold dilutions in 50mM phosphate buffer) and final volume was adjusted to 1 ml by adding 50 mM phosphate buffer. As a spectro zero 0.5 ml of 1% CMC and 0.5 ml of 50 mM phosphate buffer were used and as an enzyme control 250 µl of enzyme (2 fold dilutions) and 750 µl of 50 mM phosphate buffer was used.

The optimum pH of endoglucanase enzyme was evaluated by running DNS assay at optimum temperature (70 °C) of endoglucanase enzyme using three buffer solutions. 50 mM Na-Acetate/glycine buffer for pH 3-6, 50 mM sodium phosphate buffer for pH 7.0 and glycine-NaOH buffer for pH 8-9. One percent substrate preparation, enzyme dilution and final volume adjustment of the reaction mixture were done in respective buffers used in the enzyme assay.

Thermostability of endoglucanase enzyme in this study was evaluated by incubating enzymes at 70 °C, 80 °C and 90 °C. Enzyme incubation was carried out for 24 hrs and the activity of enzyme was measured in every 1.5 hrs for 12 hrs by running standard DNS assay. Also, the enzyme activity was measured at zero hour and 24 hours incubated samples.

4.7.5 Effect of substrate

The effect of substrate concentration on the enzyme activity of partially purified endoglucanase enzyme was determined by using different concentration of CMC and running the standard DNS assay. The assay was carried out in 1 ml reaction mixture and incubated at optimum temperature of enzyme with different CMC concentrations ranging from 0.5 mg/ml to 11 mg/ml prepared in 50 mM phosphate buffer of optimum pH 5.The value of Michael's Menten constant (*K*m) of endoglucanase was determined by linear regression plots of Lineweaver and Burk plot.

4.7.6 Glucose standard curve

Different glucose concentrations from stock of 10 mg/ml glucose were made by dissolving in 50 mM phosphate buffer (pH 6). To each 25 ml glass 1ml glucose (different concentration) and 3 ml of DNS reagent were added. As a blank 1 ml of 50 mM phosphate buffer (pH 6) was used. All tubes with glucose standards and blank were boiled for 5 minutes in a vigorously boiling water bath for color development, followed by cooling in running tap water. The mixtures in each tube were diluted by adding 20 ml of sterile MQ water and mixed completely by inverting the tubes several times so that the solution separated from bottom of tube in each inversion.

The sample dilution with 20 ml of MQ water was carried out in order to keep the absorbance of sample in linear range of spectrophotometer. Absorbance from sample was measured against the blank at 540 nm using spectrophotometer. The ΔOD_{540} values of sugar standard were plotted in Y- axis, against the calculated sugar concentration (mg/ml) in X-axis. Linear regression through the origin with the formula $y = mx + c$ is applied. Data obtained from glucose standard curve experiment is listed in Appendix 5.

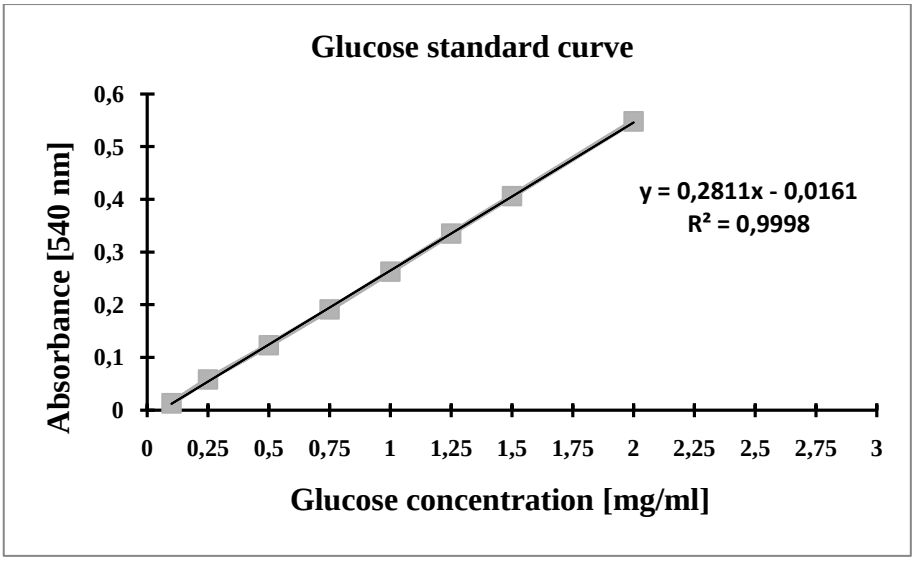


Figure 17. Standard curve of absorbance as a function of glucose concentration.

4.8 Data analysis

Endoglucanase enzyme assay for optimum temperature, pH and thermostability was carried out in 1 ml reaction mixture containing 0.5 ml substrate (1% CMC), 125 µl enzyme and 375 µl buffer. The amount of reducing sugars liberated from substrate during enzymatic reaction was determined by DNS method and taking absorbance at 540 nm. The absorbance value of samples were converted to glucose concentration (mg of glucose produced during the reaction time) with the equation (2) obtained from glucose standard curve.

$$y = 0.2811x - 0.0161 \qquad (2)$$

Once the glucose concentrations of the samples were determined they were translated into international enzyme unit i.e. CMCase (U/ml) using the equation (3). One unit of CMCase activity is defined as "µmol of glucose equivalent liberated min^{-1} under defined condition".

$$CMCase\ (Uml^{-1}\ min^{-1}) = \frac{glucose\ conc\ \ (mg/ml)}{0.18 \times 125\ (\mu l) \times 60(min)} \qquad (3)$$

Where, glucose concentration (mg/ml) is the glucose amount determined with the help of equation (1) from the assay, 0.18 corresponds to one CMCase unit when 0.18 mg min^{-1} of hydrolysis product (glucose) is formed in reaction mixture, 125 is the volume in µl of enzyme used in total reaction mixture and 60 is the incubation time for reaction mixture in minutes.

4.9 Sequence analysis and alignment

The amino acid sequence of cellulase encoding endo-1, 4-β-glucanase (EC 3.2.1.4) from thermophilic anaerobic bacteria *C. bescii* was compared to other reported cellulase gene from different bacterial species through multiple sequence alignment. Protein identification number and bacterial source for different endoglucanase are as follows: YP_002572493.1 (*C. bescii* DSM 6725), YP_001179488.1 (*C. saccharolyticus* DSM 8903), YP_003845096.1 (*C. cellulovorans* 743B), NP_622045.1 (*T. tengcongensis* MB4), and YP_003842439.1 (*C. cellulovorans* 743B). The signal sequence for the extracellular protein expression of endoglucanase from *C. bescii* was predicted by submitting the fasta format sequence to the SignalP 3.0 server (http://www.cbs.dtu.dk/services/SignalP/).

5. Results

5.1 Screening

Transformant colonies both for endoglucanase and exoglucanase were selected for screening of the plasmid with insert. Restriction analysis and PCR method were used to confirm the presence of insert in the vector. Bacterial cells harboring endoglucanase gene and cells with exoglucanase gene obtained from the screening of the transformants were streaked out in separate LA plates with 1% CMC. *E.coli* BL21 strain streaked into separate plate with similar conditions was used as a control. Congo red staining was performed to all the plates and halo zone of CMC hydrolysis was observed from bacterial cell with endoglucanase gene (Figure 18 -A) and cells with exoglucanase gene (Figure 18 -B). There was no hydrolytic zone of enzymatic activity observed in the plates with control culture.

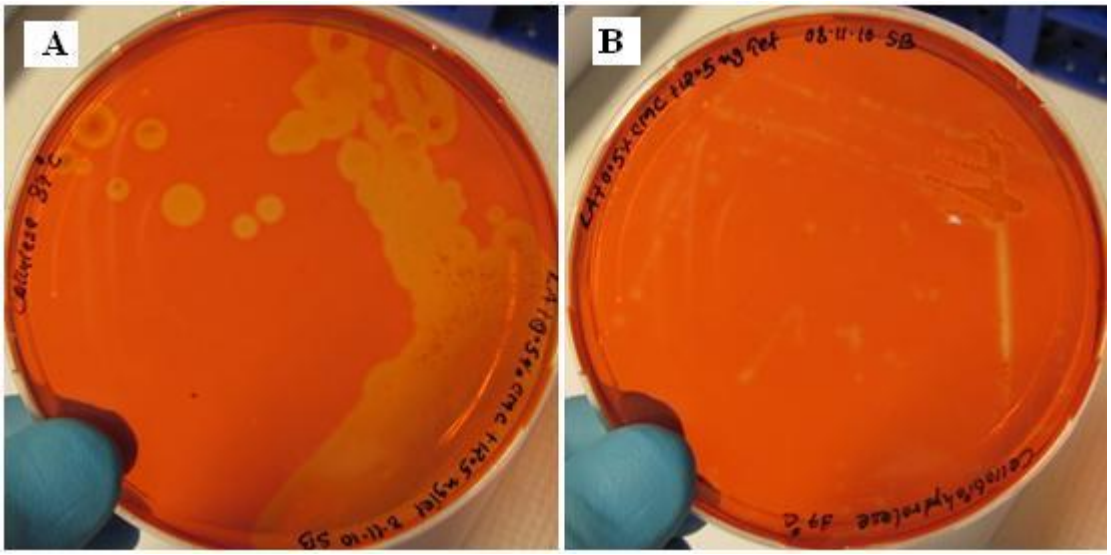


Figure 18. Plate culture for *E.coli* BL21 with pSB01-endo by Congo red assay on CMC agar plate (A) and *E.coli* BL21 with pSB02-exo by Congo red assay on CMC agar plate (B).

5.2 Preliminary enzyme assay for endoglucanase.

Endoglucanase activity of *E.coli* BL21 with pSB01-endo was carried out from culture supernatant (extracellular) and whole cell lysate (intracellular) by Congo red assay on CMC agar plate and DNS assay. Sample volume of 20 µl from both culture supernatant and cell lysate were used on CMC agar plate and Congo red staining was performed. The cell lysate of the culture showed enzyme activity as halo zones on LA plates with 1% CMC as shown in Figure 19 (A) whereas no enzyme activity was observed from culture supernatant.

The culture supernatant volume was increased from 20 µl and loaded as 50 µl, 100 µl, 150 µl and 200 µl on four separate holes punched in 1% CMC agar plate as shown in Figure 19 (B). After Congo red staining, enzyme activity as clear halo zones was noticed in all the four wells. This indicates that the extracellular secretion of endoglucanase enzyme is relatively low and most of the endoglucanase enzyme expressed remained intact inside the cells. The enzyme activity was also determined by DNS assay method from both the culture supernatant and whole cell lysate of *E.coli* BL21 with pSB01-endo on 1 % CMC. Higher enzyme activity was noticed from whole cell lysate than culture supernatant of *E.coli* BL21 with pSB01-endo (Table 11).

Table 11. Comparison of intracellular and extracellular endoglucanase activity.

Sample protein	**Average**	**Standard deviation**	**Glucose concentration**	**Cmase activity (IU ml^{-1})**
Whole cell protein	0.154	0.000	0.606	0.449
Supernatant	0.064	0.001	0.284	0.211

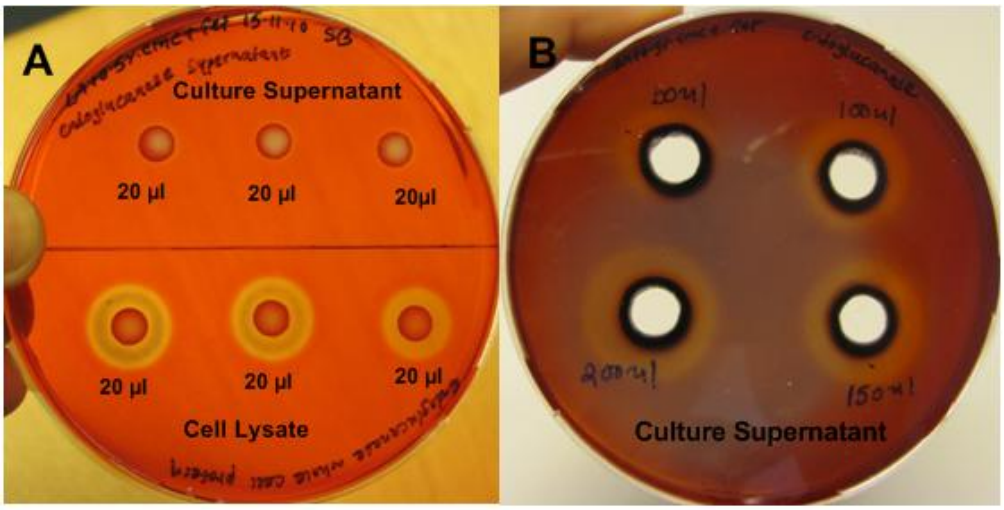


Figure 19. Endoglucanase activity study from culture supernatant and whole cell lysate of culture *E.coli* BL21with pSB01-endo in 1% CMC agar plates. 20 µl of cell lysate and culture supernatant was loaded on 1% CMC agar plate (Image A) whereas 50, 100, 150 and 200 µl of culture supernatant were used in 1% CMC agar plate (Image B) to observe the enzyme activity.

5.3 SDS PAGE for molecular mass determination

The molecular mass of endoglucanase was determined by resolving different volume of sample, control and ladder protein (Precision plus protein™ Dual Xtra standards from Bio Rad) in SDS-PAGE (Figure 20). Endoglucanase protein sample was loaded as 30 µl in lane 1 and 20 µl in lane 3, whereas control protein sample of 30 µl in lane 2 and 20 µl in lane 4. Protein ladder of 5 µl was loaded in three different lanes represented by ladder in the Figure 20. The SDS-PAGE gel image indicates that almost all protein bands present in endoglucanase protein sample are present in control protein sample except one extra band of nearly 75 kDa. This extra band indicates the presence of endoglucanase protein which has a predicted protein size of ~85 kDa. The extra band of 75 kDa is also observed in lane 3 with endoglucanase protein sample (20 µl) and not in lane 4 with control protein sample (20µl). Note: the amount of protein concentration in control and endoglucanase sample were not quantified as a result of which the protein bands showed different intensity in the gel image after resolving in SDS-PAGE which suggests different protein concentration between the control and endoglucanase samples in the gel image.

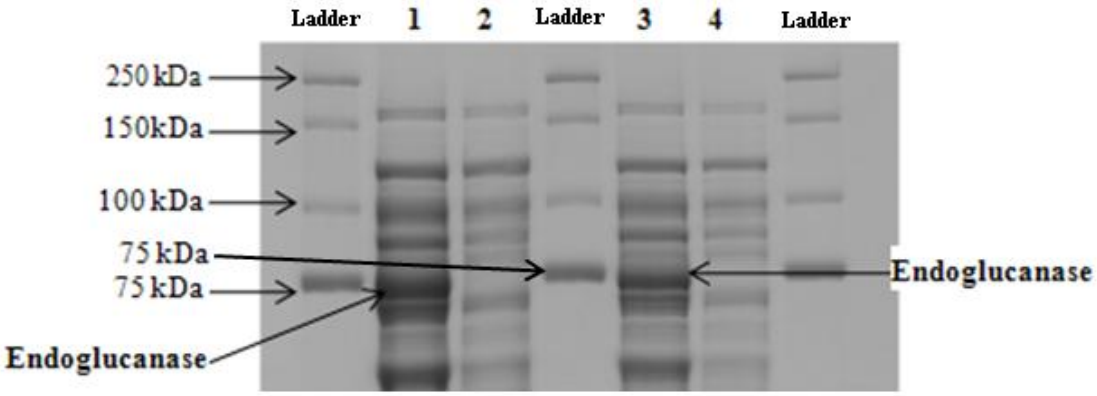


Figure 20. SDS-PAGE analysis of endoglucanase from *E.coli* BL21 with plasmid pSB01-endo and control protein *E.coli* BL21. Ladder denotes 5 µl of protein ladder in each lane.

5.4 Optimum temperature determination

The temperature dependence of *E.coli* produced endoglucanase activity on 1% CMC was determined at pH 5 by measuring enzyme activity at various temperatures ranging from 40 °C to 90 °C by running standard DNS assay. Graph plotted with the results obtained from DNS assay of endoglucanase activity at different temperature (Appendix 6) indicates that the enzyme is active at broad temperature range of 40 °C to 75 °C. A gradual increase in enzyme activity was observed with an increase in temperature till it reached its optimal temperature range at 70 °C. Beyond its optimal temperature range, the enzyme activity started to decline and minimum activity was observed at 90 °C (Figure 21).

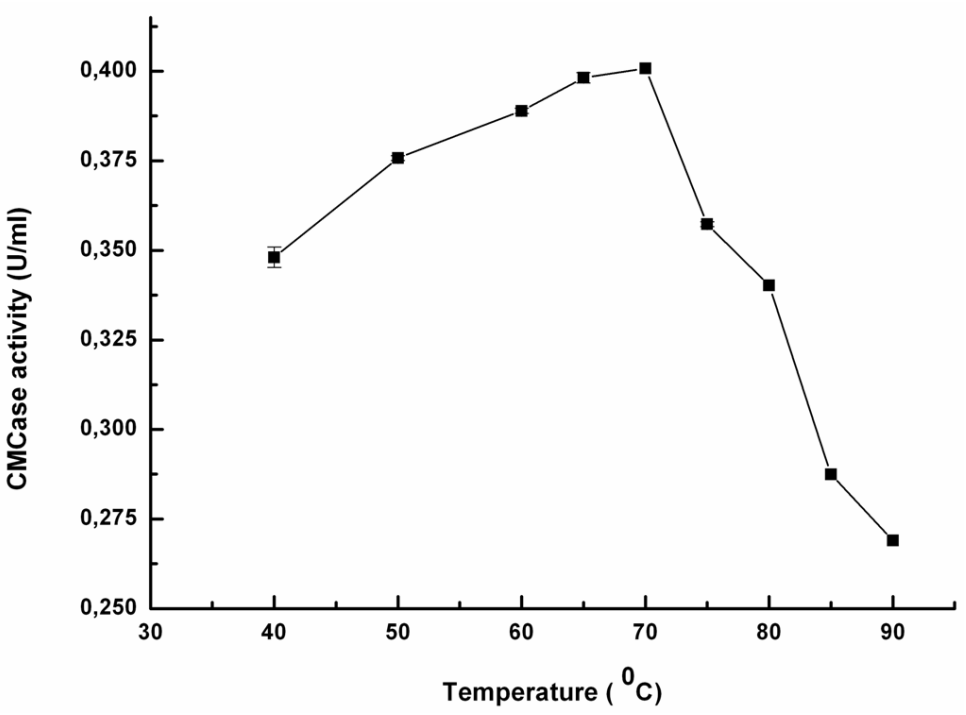


Figure 21. Effect of temperature on recombinant endoglucanase activity. Enzyme activity was assayed at pH 5 for 60 min at the indicated temperature range.

5.5 Optimum pH

The effect of pH on activity of recombinant endoglucanase enzyme was determined at 70 °C using various buffers ranging from pH 3 to pH 10. The graph plotted from the results obtained from enzyme assay at different pH range (Appendix 7), suggests that the enzyme is active at both acidic and basic pH. The maximum activity of the enzyme was observed at pH 5 and a negligible activity was observed at pH 3. The decrease in enzyme activity was observed with gradual increase in pH values starting from the optimal pH 5 (Figure 22).

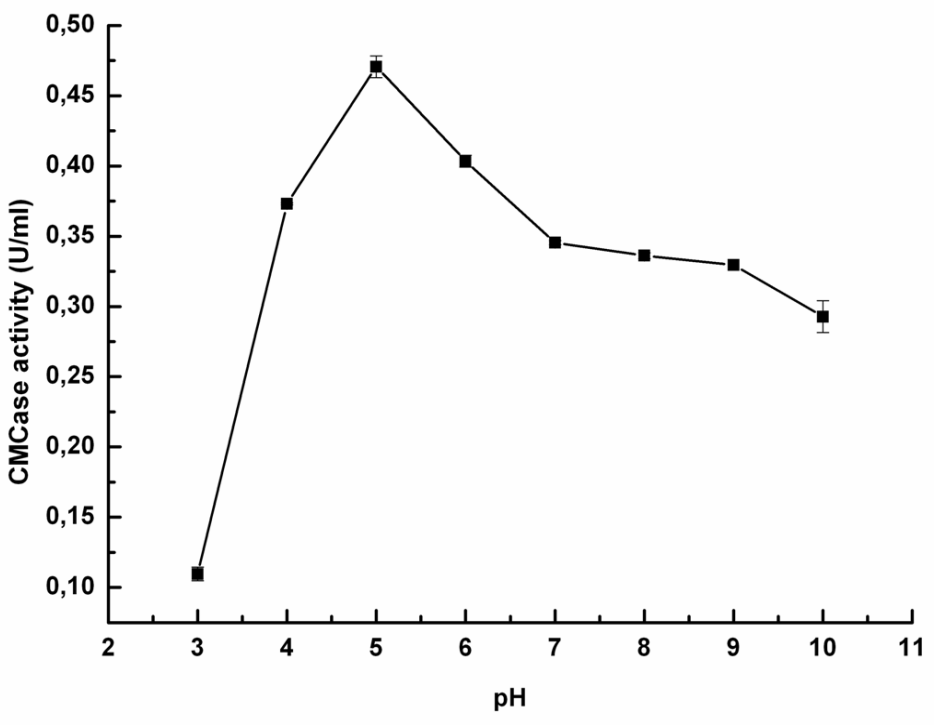


Figure 22. The optimum pH of recombinant endoglucanase determined by DNS assay at 70 ℃ using 1 % CMC as substrate at different pH.

5.6 Thermostability test

The effect of temperature on the stability of *E.coli* produced recombinant endoglucanase enzyme was examined by measuring the decrease in enzyme activity after incubation at high temperature. The test was carried out by continuous incubation of the enzyme at 70 ℃, 80 ℃ and 90 ℃.

In Figure 23, the activity of endoglucanase decreased drastically after 1.5 hours incubation at 90 ℃ and the activity was almost constant for extended incubation till 24 hours. The enzyme activity at 80 ℃ did not decrease drastically as it was observed from samples at 90 ℃. A slight decrement of enzyme activity was noticed until 6 hours of incubation, however, increased loss of activity was observed from samples incubated longer than 6 hours. Enzyme samples incubated at 70 ℃ did not show any loss in enzyme activity throughout the incubation time even after the end of 24 hours. This indicates that the enzyme is highly stable at its optimum temperature and can retain maximum activity even after 24 hours incubation. Experimental data obtained during this test is listed in Appendix 8.

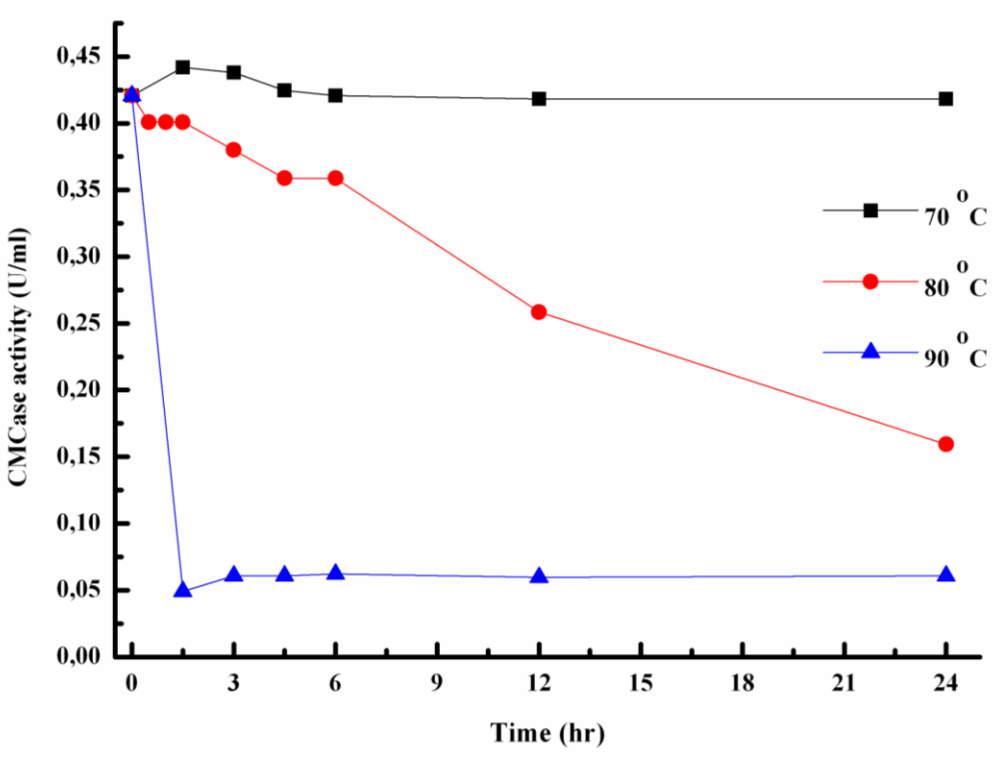


Figure 23. Effect of temperature on recombinant endoglucanase enzyme activity and stability.

5.7 Effect of substrate

The Michaelis- Menten constant (K_m) of endoglucanase enzyme was calculated by linear regression plots of Lineweaver-Burk as shown in Figure 24. An inverse of substrate concentration (1/S) was plotted along the X-axis and inverse of enzyme activity also known as enzyme velocity along the Y-axis. The maximum velocity of enzyme (V_{max}) was determined as 1.069 U min^{-1} while K_m was 7.288 mg/ml and the specificity constant [V_{max}/ K_m] was 0.14. All the data obtained during this experiment is mentioned in Appendix 9 and 10.

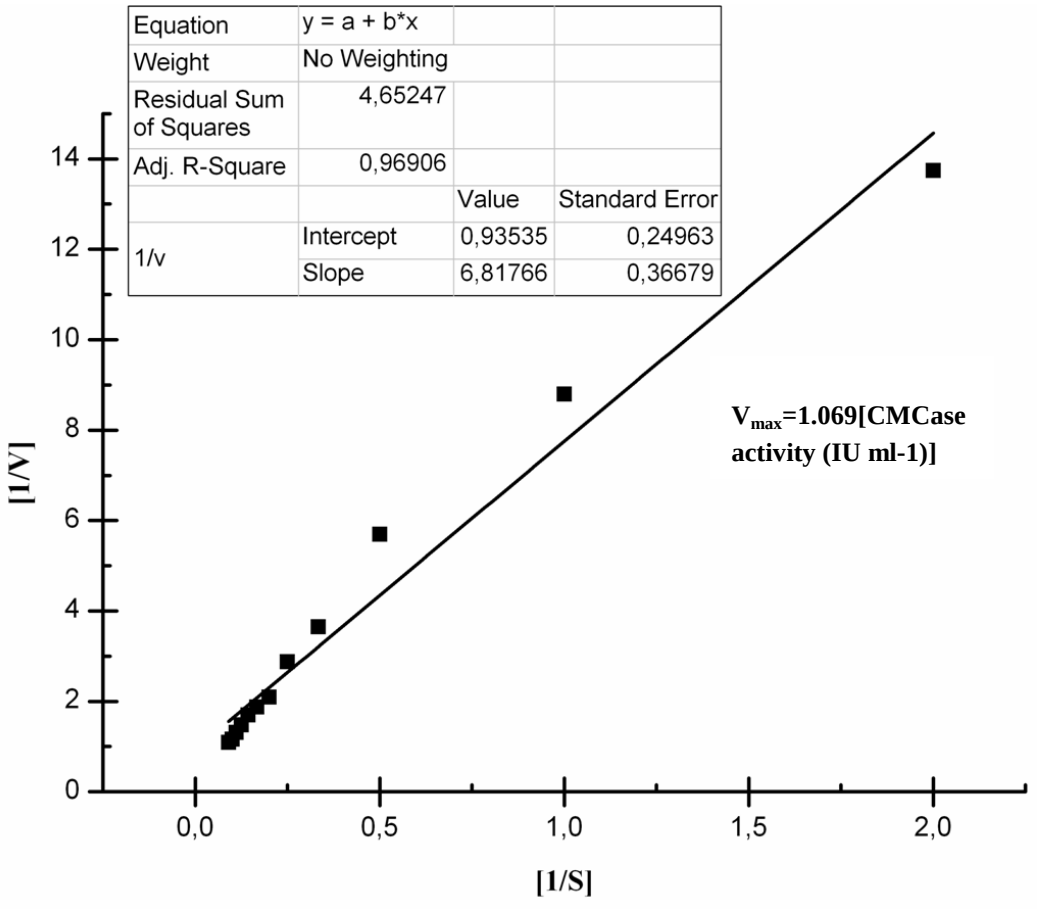


Figure 24. Double reciprocal plot for determination of Michaeli's- Menten kinetic constants (V_{max} and K_m) of endoglucanase.

5.8 Sequence analysis and homology of endoglucanase

The endoglucanase gene of *C. bescii* contains 2268 base pairs encoding a polypeptide of 755 amino acids with a calculated molecular weight of 82.514 kDa. The first 30 amino acids of endoglucanase show the typical feature of a prokaryotic signal peptide at its N terminus, with a calculated molecular weight of 3446 Da. The maximum cleavage site probability in the signal sequence was found to be 0.999 between the 29th and 30th amino acid residue i.e. between the VFA-QS as shown in Figure 25.

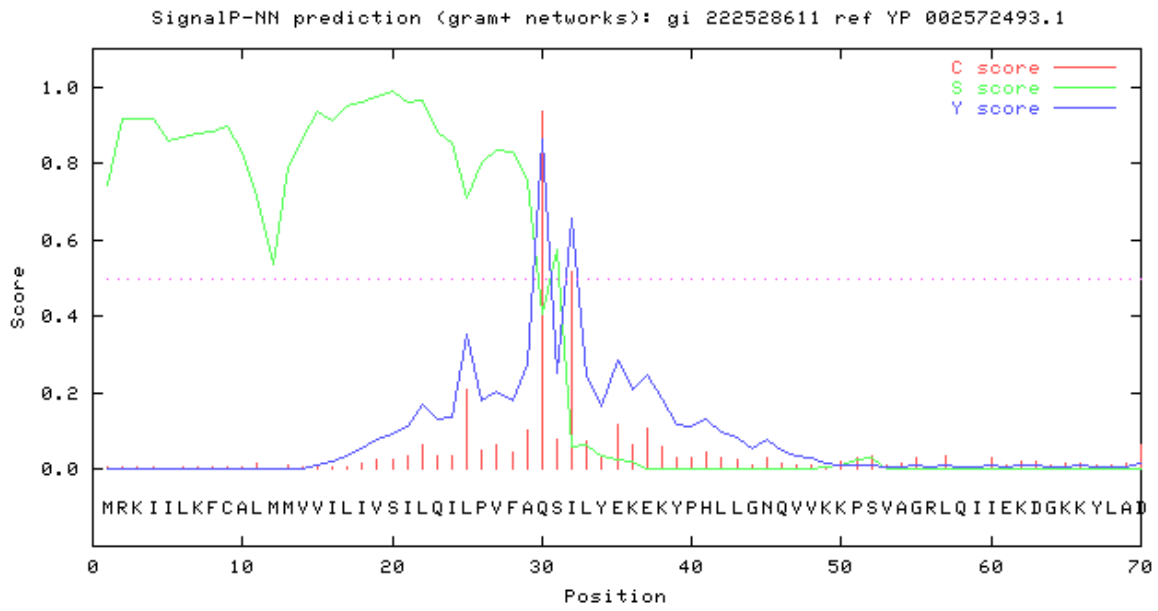


Figure 25. Signal sequence prediction of endoglucanase gene from *C. bescii* and determination of cleavage probability.

In its sequence, endoglucanase harbors a catalytic domain of glycoside hydrolases family 2 (GH2; residues ~ 60 to ~ 330) followed by an incomplete family 17_28 carbohydrate binding module (CBM; from residues ~ 352). Sequence alignment results showed that endoglucanase (*C. bescii*) shares the closest similarity (identity 71 %) to endoglucanase of *T. tengcongensis* MB4 and 65 % identity to endoglucanase of *C. saccharolyticus* (Table 12).

Table 12. Sequence similarity result of endoglucanases from sequence alignment analysis.

	YP_003842439.1	YP_003845096.1	NP_622045.1	YP_001179488.1	YP_002572493.1
YP_003842439.1		55	51	32	32
YP_003845096.1			56	42	41
NP_622045.1				70	71
YP_001179488.1					65
YP_002572493.1					

Conserved amino acid sequence of various lengths was observed by multiple sequence alignment of endoglucanase from *C. bescii* and other selected endoglucanase. Highly conserved amino acids were observed in position 76, 108, 144, 155, 183 and 190 of the endoglucanase sequences (Figure 26).

Figure 26. Multiple sequence alignment of endoglucanase (*C. bescii*) with other reported endoglucanase from different bacterial species. Amino acid residues in yellow color indicate the strictly conserved residue and the conservatively substituted residues are highlighted with blue color.

Phylogenetic tree analysis of all the endoglucanase used in multiple sequence alignment was carried out. According to the analysis of the phylogenetic tree endoglucanase of *C.bescii* (YP_002572493.1) is very close to the endoglucanase of *C.saccharolyticus* (YP_001179488.1) however the sequence similarity was comparatively lower than the endoglucanase of *T. tengcongensis* MB4 (NP_622045.1) as shown in Figure 26.

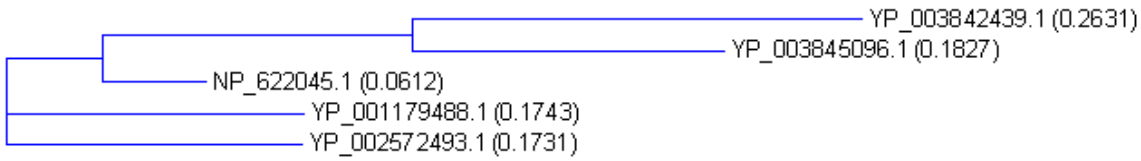


Figure 27. Phylogenetic tree using amino acid sequences of endo-β-1,4 glucanase homologs from YP_002572493.1 (*C.bescii* DSM 6725), YP_001179488.1 (*C. saccharolyticus* DSM 8903), YP_003845096.1(*C.cellulovorans* 743B), NP_622045.1 (*T. tengcongensis* MB4), and YP_003842439.1 (*C. cellulovorans* 743B).

6. Discussion

6.1 Experimental discussion

This laboratory work describes the expression, purification and enzymatic characterization of novel thermostable cellulase gene (GeneID: "7406935") which encodes an endo-1, 4-β-D-glucanase (EC_number="3.2.1.4") from thermophilic bacterium *C. bescii*. It has a coding sequence corresponding to a calculated polypeptide of 82.154 kDa, including a typical prokaryotic signal peptide of 30 amino acids for extracellular protein expression across the cytoplasmic membrane. This cellulase (endoglucanase) gene was reported from the genome sequence study of *C. bescii* and annotated in gene bank (Kataeva et al. 2009). To the best of our knowledge, there have been no studies reported on the characterization of this cellulase.

At first, *E.coli* BL21 with plasmid pSB01-endo and only *E.coli* BL21 as control were grown in LB medium supplemented with 0.5 % glucose and protein expression was induced in presence of 0.1 mM IPTG. The whole cell lysate and culture supernatant of both the cultures with plasmid pSB01-endo and without plasmid (control) were used to perform DNS test for endoglucanase enzyme activity in presence of 1% CMC. The initial data obtained and analysis showed lower expression of endoglucanase gene as a result of which less enzyme activity was found from both the cell lysate and cell supernatant (data not shown). Apart from this, enzyme activity was also obtained from the cell lysate and cell supernatant of the control culture without endoglucanase gene. Later it was found that the addition of 0.5 % glucose in the culture media as extra carbon source for better cell growth was responsible for the false enzyme activity. Since DNS reagent used for endoglucanase assay binds with the left over sugar molecules in the medium (Ghose TK 1987; Miller 1959). Thus, the growth medium for cell growth and protein expression was shifted to nutrient enriched (2X YT) medium instead of LB medium and induction of protein expression was carried out using in 0.4 mM IPTG which is recommended concentration for full induction of protein in expression vector carrying “plain” T7 lac promoter (pET system manual 10th edition, Novagen).

Cell cultures were grown in 2X YT medium for protein production. Culture supernatant and whole cell lysate after heat treatment to remove all the heat-labile proteins (Hung et al. 2011) was used for the preliminary enzyme assay using Congo red and DNS method (Figure 19 and Table 11). Higher enzymatic activity was observed from the whole cell lysate of the partially purified protein sample in comparison to the culture supernatant sample. This suggests that endoglucanase expressed was not exported extracellularly since

higher enzyme activity was observed from the whole cell lysate protein sample. Lower enzyme activity from the culture supernatant can arise due to the dilution of extracellular excreted enzyme in the culture medium and lack of enzyme concentration. The reduced enzyme concentration might have arise due to the signal peptide present in the endoglucanase gene, which is from gram positive bacteria and probably was not optimal for the extracellular protein expression in *E.coli*. Thus, the heat treated whole cell lysate was used to characterize the endoglucanase from *C. bescii*.

Molecular weight determination of the endoglucanase was carried out by SDS-PAGE method by resolving the protein sample in 10% resolving gel. A protein with molecular weight around 75 kDa was observed from the cell lysate of *E.coli* culture harboring the plasmid with endoglucanase gene (Figure 20). The extra band was not observed from the cell lysate of control culture (*E.coli* BL21). The calculated molecular mass of the endoglucanase is 81.264 kDA; however the molecular mass on the SDS-PAGE appeared around 75 kDA. This difference in molecular mass might have resulted due to the expression of recombinant protein in *E.coli* which has proteolytic degradation ability to render the isolation of full length gene product impossible. This could also arise due to the deficiency in glycosylation of the expressed protein by the host cell or high sensitivity to the protelolytic enzymes produced by heterologous host (Zverlov et al. 1998).

As shown in Figure 21, determination of optimum temperature of endoglucanase from *C. bescii* was examined after incubation at various temperatures ranging from 40 ℃ to 90 ℃ for 60 min at pH 5. Higher hydrolytic activity against CMC was observed at temperatures between 40 ℃ to 75 ℃ with maximum activity at 70 ℃. Lowest enzyme activity was observed at temperature 85 ℃ and 95 ℃ in comparison to other temperatures which may be attributed to the enzyme denaturation with increase in temperature. Thus, the enzyme is considered as thermostable with optimum temperature at 70 ℃. Compared to endoglucanase isolated from *Bacillus* sp-A8-8 which has the optimal temperature at 60 ℃ (Jung et al. 2007), the endoglucanase from *C. bescii* has highest activity at 70 ℃. The optimal temperature of thermostable cellulase isolated from *B. subtilis* strain I15 was 65 ℃ (Yang et al. 2010a) which is relatively lower than the endoglucanase from *C. bescii* found during this study. Endoglucanase enzyme isolated from mesophilic bacteria like *E.coli* has optimal activity at a temperature of 40 ℃ (Park et al. 1999) which is relatively lower than the endoglucanase isolated from *C. bescii* with optimal activity at 70 ℃ determined during this study.

The optimum activity of endoglucanase from *T. tengcongensis* MB4 was 75 °C (Liang et al. 2011) compared to optimal growth temperature (75 °C). However, the maximum enzyme activity of endoglucanase from *C. bescii* was found at 70 °C although the bacteria had similar optimum growth temperature to *T. tengcongensis* MB4. This observation suggests that the enzyme from thermophilic bacteria has higher optimum enzyme activity.

Measurement of pH profile of endoglucanase was carried out at temperature 70 °C. The enzyme was found to be active at wide pH range. The lower and the upper value of the enzyme activity were determined to be pH 4 and pH 10. However, the maximum activity was noticed at pH 5 (Figure 22). Similar results were observed in *C. bescii* and *C. obsidiansis* in which the cellulolytic enzymes hydrolyzed CMC optimally at pH 5 (Lochner et al. 2011). The experimental data obtained from the pH experiment suggest that the endoglucanase from *C. bescii* is highly stable over a wide pH range with activity from acidic pH 5.0 to alkaline pH 10.

The broad pH range of endoglucanase enzyme activity from *C. bescii* is similar to most of the thermostable endoglucanase reported till date. Thermostable endoglucanase isolated from *Bacillus* sp. A8-8 has a broad pH range (pH 3- pH 10) of enzyme activity (Jung et al. 2007). Similar finding has been reported from the halo-tolerant endoglucanase of *T. tengcongensis* MB4 which has enzyme activity over a wide pH range (pH 5- pH 9) at 75 °C (Liang et al. 2011). Thus, the highly stable and pH tolerant endoglucanase from *C. bescii* can be an ideal candidate for industrial application.

Thermostability of the enzyme was monitored by continuous incubation of the enzyme at temperature 70 °C, 80 °C and 90 °C for almost 24 hours and measuring the residual enzyme activity in every 90 minutes (Figure 23). The enzyme was highly stable at 70 °C during the 24 hours of incubation retaining its 100 % activity. However, the gradual loss of enzyme activity was noticed at temperature 80 °C with increase of incubation time. Temperature at 90 °C sharply decreased the enzyme activity after one and half hours incubation. In comparison to the endoglucanase reported by Liang et al. 2011 from *T. tengcongensis* MB4 with optimum temperature 75 °C the enzyme was able to maintain only 80% of maximum activity after incubation at 60 °C for 24 hrs.

Most of the cellulase enzymes studied from *Bacillus* species has thermostability for a short period of time. Thermostable cellulase isolated from *B. subtilis* strain I15 was found to be stable at 65 °C for 2 hours with retaining its 90 % cellulase activity (Yang et al. 2010a). Similar, result was observed from the thermostable endoglucanase from *Bacillus* sp. A8-8 which retained over 70 % of its original activity after incubation at 80 °C for 2 hours (Jung et al. 2007). Compared to these thermostable endoglucanases isolated from *Bacillus* species, the endoglucanase characterized from *C. bescii* during this study showed

better thermostability at 70 °C and 80 °C for a longer period of time. It retained more than 90 % of its activity at 70 °C for 24 hrs of incubation and more than 70 % of the CMCase was maintained at 80 °C after incubation for 6 hours (Figure 23).

Thermostable endoglucanase enzyme isolated from *C. thermophilum* and expressed in *E.coli* showed the optimum temperature at 75 °C, however the enzyme was thermostable at 60 °C for 2 hours of incubation retaining its maximum activity above 90% (Schwarz et al. 1986). Compared to the endoglucanase from *C. thermophilum*, the thermostability of *C. bescii* endoglucanase shows better thermostability for a longer time period (24 hours) at its optimum temperature 70 °C. The unique property of the endoglucanase to hydrolyze cellulase at 70 °C to 80 °C, higher thermostability and wide pH range makes this enzyme as a potential candidate for industrial application like cellulose hydrolysis and biopolishing of cotton products.

The Michaelis-Menten constant (K_m) and maximum velocity (V_{max}) of *C. bescii* endoglucanase was found lower in comparison to other reported endoglucanase. This could be due to the lower yield of enzyme from *E.coli* transformants and the lack of proper technique to concentrate and purify the enzyme. In order to achieve more efficient extracellular production of the endoglucanase, gram positive host cells like *Bacillus* (Ando et al. 2002) and *S. lividans* (Vrancken et al. 2010) could be employed. In fact, *C. bescii* is a gram positive bacteria and the signal sequence of endoglucanase can be better recognized and utilized by the gram positive expression host.

Widespread groups of enzymes belong to the glycoside hydrolases (GH, EC 3.2.1-) that hydrolyze glycosidic bond present in the carbohydrate backbone. Based on amino acid sequence similarity or folding similarity they have been classified into 118 families (Liang et al. 2011). Endoglucanase (EC 3.2.1.4) from *C. bescii* contained an open reading frame (ORF) which starts with an ATG start codon and terminates with a TAG stop codon. The ORF of endoglucanase consists of 2268 nucleotides encoding a protein of 755 amino acids. Sequence analysis shows that the amino acid sequence consist a glycosyl hydrolases family 2 domain and a carbohydrate binding domain. Sequence alignment results showed that the *C. bescii* endoglucanase showed highest homology of (71%) to the cel5A of *T. tengcongensis* MB4. Similar homology result was reported by Liang et al. 2011. Second highest homology of 65 % was observed with the endoglucanase from *C. saccharolyticus*. The predicted amino acid sequence showed low homology (only 32 % and 41 %) with the cellulase gene from *C. cellulovorans* 743B. Sequence alignment result among the selected endoglucanase from various bacterial species showed that they contain highly conserved amino acid residues at different position of the sequences. Glutamic acid residues in the conserved sequence (E186, E295) acts as a proton donor and nucleophile for hydrolyzing

β-1,4-glycosidic bond are strictly conserved within this family and identified as a catalytic center (Posta et al. 2004).

Amino acid sequence analysis of endoglucanase from *C. bescii* indicates that it contains 70 Isoleucine (I) and 50 Valine (V) residues which cover as the 1st and 5th highest amino acid residues in the total polypeptide chain. Also, the conserved amino acid residues in sequence alignment show higher presence of Isoleucine and Valine. Dominance of two strong hydrophobic amino acids Isoleucine and Valine which contains two strong hydrophobic substituents compared to other amino acids in the polypeptide chain of endoglucanase can be suggested as one of the main factor influencing the thermostability of enzyme. In fact, hydrophobic interaction has been proposed to play a crucial role in the stabilization of enzymes' structures at high temperature (Li et al. 2008).

6.2 Suggestions for future research

I. Characterization of cellobiohydrolase gene from *C. bescii* that has been cloned in vector pVKK81 named as pSB-02.

II. New sets of forward primer can be designed for both the endoglucanase and cellobiohydrolase gene introducing 6xHis or poly-His tag. Expressed His-tagged proteins can be purified and detected easily.

III. Commercial vectors like pET vectors from Novagen can be used for recombinant protein expression instead of pVKK81.

IV. Expression of recombinant plasmid with endoglucanase and exoglucanase gene can be performed in gram positive host cells like *Bacillus* species, since the cellulase gene are obtained from gram positive bacteria and signal sequence of this gene can be well recognized by gram positive host cells for extracellular protein expression.

7. Conclusion

Endoglucanase and exoglucanase (cellobiohydrolase) genes from thermophilic anaerobic bacteria *C. bescii* were successfully cloned and expressed in heterologous host *E.coli*. Analysis of the enzyme activity was performed from the crude protein extract as well as from the culture supernatant. DNS assay and Congo red assay confirmed the higher enzyme activity from the crude protein extract when compared to culture supernatant. Enzyme kinetics was studied using the crude protein extract. The optimum pH and temperature of the enzyme was determined. Moreover, the activity of the enzyme at different pH (3-10), temperature (40 °C - 90 °C) and with different substrate concentration (0.5 mg/ml - 11 mg/ml) was studied during this research.

The enzyme was found to be highly stable over a broad range of temperature from 40 °C to 80 °C and pH range from 4 to 10. The optimum pH and temperature for the highest activity of the enzyme was found to be at pH 5 and 70 °C respectively. The enzyme was found to be highly thermostable at 70 °C and retained its maximum activity even after 24 hours of incubation. Comparatively, thermostability of endoglucanase at 80 °C was found lower than at 70 °C. Gradual decrease in enzyme activity along with increase in incubation time was observed at 80 °C. On the basis of sequence alignment study, endoglucanase from *C. bescii* has maximum similarity of 71% with the endoglucanase of *T. tengcongensis* MB4 and 65 % similarity with the endoglucanase of *C. saccharolyticus*. Also, highly conserved amino acid regions were found with other reported endoglucanase through alignment study.

The unique property of this characterized endoglucanase with broad pH range, high optimum temperature and thermostability makes a potential candidate among other reported cellulase for industrial applications. It can be used in biopolishing of cotton products, textile industry and cellulose hydrolysis in synergism with other cellulolytic enzyme in industrial applications.

References

Agbor VB, Cicek N, Sparling R, Berlin A, Levin DB (2011) Biomass pretreatment: Fundamentals toward application. Biotechnol Adv In Press, Uncorrected Proof

Ahmad AL, Yasin NHM, Derek CJC, Lim JK (2011) Microalgae as a sustainable energy source for biodiesel production: A review. Renewable and Sustainable Energy Reviews Volume 15:584-593

Ajanovic A (2011) Biofuels versus food production: Does biofuels production increase food prices? Energy 36:2070-2076

Ando S, Ishida H, Kosugi Y, Ishikawa K (2002) Hyperthermostable endoglucanase from *Pyrococcus horikoshii*. Appl Environ Microbiol 68:430-433

Badal S (2004) Lignocellulose Biodegradation and Applications in Biotechnology. American Chemical Society, Washington, DC, pp 2-34

Badel S, Laroche C, Gardarin C, Petit E, Bernardi T, Michaud P (2011) A new method to screen polysaccharide cleavage enzymes. Enzyme Microb Technol 48:248-252

Bastawde KB (1992) Xylan structure, microbial xylanases, and their mode of action. World Journal of Microbiology and Biotechnology 8:353-368

Bayer E, Shoham Y, Lamed R (2006) Cellulose-decomposing Bacteria and Their Enzyme Systems. The Prokaryotes 2:578-617

Blumer-Schuette SE, Lewis DL, Kelly RM (2010) Phylogenetic, microbiological, and glycoside hydrolase diversities within the extremely thermophilic, plant biomass-degrading genus *Caldicellulosiruptor*. Appl Environ Microbiol 76:8084-8092

Bredholt S, Mathrani IM, Ahring BK (1995) Extremely thermophilic cellulolytic anaerobes from icelandic hot springs. Antonie van Leeuwenhoek, International Journal of General and Molecular Microbiology 68:263-271

Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B (2009) The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. Nucleic Acids Res 37:D233-8

Cao Y, Tan H (2002) Effects of cellulase on the modification of cellulose. Carbohydr Res 337:1291-1296

Carere CR, Sparling R, Cicek N, Levin DB (2008) Third generation biofuels via direct cellulose fermentation. Int J Mol Sci 9:1342-1360

Chandel AK, Singh OV (2011) Weedy lignocellulosic feedstock and microbial metabolic engineering: advancing the generation of 'Biofuel'. Appl Microbiol Biotechnol 89:1289-1303

Charlier D, Droogmans L (2005) Microbial Life at high temperature, the challenges, the strategies. Cellular and Molecular Life Sciences 62:2974-2984

Dam P, Kataeva I, Yang S, Zhou F, Yin Y, Chou W, Poole FL, Westpheling J, Hettich R, Giannone R, Lewis DL, Kelly R, Gilbert HJ, Henrissat B, Xu Y, Adams MWW (2011) Insights into plant biomass conversion from the genome of the anaerobic thermophilic bacterium *Caldicellulosiruptor bescii* DSM 6725. Nucleic Acids Res 39:3240-3254

Dashtban M, Schraft H, Qin W (2009) Fungal bioconversion of lignocellulosic residues; opportunities & perspectives. Int J Biol Sci 5:578-595

Dashtban M, Maki M, Leung KT, Mao C, Qin W (2010) Cellulase activities in biomass conversion: Measurement methods and comparison. Crit Rev Biotechnol 30:302-309

DeBoy RT, Mongodin EF, Fouts DE, Tailford LE, Khouri H, Emerson JB, Mohamoud Y, Watkins K, Henrissat B, Gilbert HJ, Nelson KE (2008) Insights into plant cell wall degradation from the genome sequence of the soil bacterium *Cellvibrio japonicus*. J Bacteriol 190:5455-5463

Dellomonaco C, Fava F, Gonzalez R (2010) The path to next generation biofuels: successes and challenges in the era of synthetic biology. Microb Cell Fact 9:3

Demain AL, Newcomb M, Wu JH (2005) Cellulase, clostridia, and ethanol. Microbiol Mol Biol Rev 69:124-154

Demirbas A (2009a) Biofuels securing the planet’s future energy needs. Energy Conversion and Management 50:2239-2249

Demirbas A (2009b) Political, economic and environmental impacts of biofuels: A review. Applied Energy 86: S108–S117

Ding SY, Xu Q, Crowley M, Zeng Y, Nimlos M, Lamed R, Bayer EA, Himmel ME (2008) A biophysical perspective on the cellulosome: new opportunities for biomass conversion. Curr Opin Biotechnol 19:218-227

Doi RH, Kosugi A (2004) Cellulosomes: plant-cell-wall-degrading enzyme complexes. Nat Rev Microbiol 2:541-551

Ghose TK (1987) Measurement of cellulase activities. Pure Appl Chem 59:257-268

Gomes J, Steiner W (2004) The biocatalytic potential of extremophiles and extremozymes. Food Technol. Biotechnol 42:223–235

Gong Y, Jiang M (2011) Biodiesel production with microalgae as feedstock: from strains to biodiesel. Biotechnology Letters 33:1269-1284

Gowen CM, Fong SS (2010) Exploring biodiversity for cellulosic biofuel production. Chem Biodivers 7:1086-1097

Haki GD, Rakshit SK (2003) Developments in industrially important thermostable enzymes: a review. Bioresour Technol 89:17-34

Harry JG, Geoffrey PH (1993) Bacterial cellulases and xylanases. Journal of General Microbiology 139:187-194

Henry RJ (2010) Evaluation of plant biomass resources available for replacement of fossil oil. Plant Biotechnology Journal 8:288-293

Hill J, Polasky S, Nelson E, Tilman D, Huo H, Ludwig L, Neumann J, Zheng H, Bonta D (2009) Climate change and health costs of air emissions from biofuels and gasoline. Proc Natl Acad Sci U S A 106:2077-2082

Howard RL, Abotsi E, Van Rensburg ELJ, Howard S (2003) Lignocellulose biotechnology: Issues of bioconversion and enzyme production. African Journal of Biotechnology 2:702-733

Hung KS, Liu SM, Fang TY, Tzou WS, Lin FP, Sun KH, Tang SJ (2011) Characterization of a salt-tolerant xylanase from *Thermoanaerobacterium saccharolyticum* NTOU1. Biotechnol Lett 33:1441-1447

Jeffries TW (1994) Biodegradation of lignin and hemicelluloses. In: Anonymous Biochemistry of Microbial Degradation. Kluwer Academic Publishers, Dordrecht, Netherlands, pp 233-277

Jørgensen H, Kristensen JB, Felby C (2007) Enzymatic conversion of lignocellulose into fermentable sugars: challenges and opportunities. Biofuels, Bioprod. Bioref 1:119-134

Jung Y, Yoo J, Lee Y, Park I, Kim S, Lee S, Yasuda M, Chung S, Choi Y (2007) Purification and characterization of thermostable β-1,3-1,4 glucanase from *Bacillus* sp. A8-8. Biotechnology and Bioprocess Engineering 12: 265-270

Kataeva IA, Yang SJ, Dam P, Poole FL,2nd, Yin Y, Zhou F, Chou WC, Xu Y, Goodwin L, Sims DR, Detter JC, Hauser LJ, Westpheling J, Adams MW (2009) Genome sequence of the anaerobic, thermophilic, and cellulolytic bacterium "*Anaerocellum thermophilum*" DSM 6725. J Bacteriol 191:3760-3761

Khraisheh M, Li A (2010) Bio-ethanol from Municipal Solid Waste (MSW): The Environmental Impact Assessment. In: Anonymous Proceedings of the 2nd Annual Gas Processing Symposium. Elsevier, Amsterdam, pp 69-76

Kumar R, Singh S, Singh OV (2008) Bioconversion of lignocellulosic biomass: biochemical and molecular perspectives. J Ind Microbiol Biotechnol 35:377-391

Lee J (1997) Biological conversion of lignocellulosic biomass to ethanol. J Biotechnol 56:1-24

Lehninger A, Nelson DL, Cox MM (2008) Lehninger Principles of Biochemistry. W. H. Freeman,

Li W, Zhang W-, Yang M-, Chen Y- (2008) Cloning of the thermostable cellulase gene from newly isolated *Bacillus subtilis* and its expression in *Escherichia coli*. Mol Biotechnol 40:195-201

Liang C, Xue Y, Fioroni M, Rodríguez-Ropero F, Zhou C, Schwaneberg U, Ma Y (2011) Cloning and characterization of a thermostable and halo-tolerant endoglucanase from *Thermoanaerobacter tengcongensis* MB4. Appl Microbiol Biotechnol 89:315-326

Lochner A, Giannone RJ, Rodriguez M,Jr, Shah MB, Mielenz JR, Keller M, Antranikian G, Graham DE, Hettich RL (2011) Use of label-free quantitative proteomics to distinguish the secreted cellulolytic systems of *Caldicellulosiruptor bescii* and *Caldicellulosiruptor obsidiansis*. Appl Environ Microbiol 77:4042-4054

Lynd LR, Weimer PJ, van Zyl WH, Pretorius IS (2002) Microbial cellulose utilization: fundamentals and biotechnology. Microbiol Mol Biol Rev 66:506-77

Maki M, Leung KT, Qin W (2009) The prospects of cellulase-producing bacteria for the bioconversion of lignocellulosic biomass. Int J Biol Sci 5:500-516

Marchant R, Banat IM, Rahman TJ, Berzano M (2002) The frequency and characteristics of highly thermophilic bacteria in cool soil environments. Environ Microbiol 4:595-602

Miller GL (1959) Use of Dinitrosalicylic Acid Reagent for Determination of Reducing Sugar. Anal Chem 31:426-428

Minic Z, Jouanin L (2006) Plant glycoside hydrolases involved in cell wall polysaccharide degradation. Plant Physiol Biochem 44:435-449

Moreira LR, Filho EX (2008) An overview of mannan structure and mannan-degrading enzyme systems. Appl Microbiol Biotechnol 79:165-178

Naik SN, Goud VV, Rout PK, Dalai AK (2010) Production of first and second generation biofuels: A comprehensive review. 14:578-597

Nigam PS, Singh A (2011) Production of liquid biofuels from renewable resources. Progress in Energy and Combustion Science 37:52-68

Park YW, Yun HD (1999) Cloning of the *Escherichia coli* endo-1,4-D-glucanase gene and identification of its product. Mol Gen Genet 261:236-241

Patil V, Tran K, Giselrod H (2008) Towards Sustainable Production of Biofuels from Microalgae. Int J Mol Sci 9:1188-1195

Posta K, Béki E, Wilson DB, Kukolya J, Hornok L (2004) Cloning, characterization and phylogenetic relationships of cel5B, a new endoglucanase encoding gene from Thermobifida fusca. J Basic Microbiol 44:383-399

Ruane J, Sonnino A, Agostini A (2010) Bioenergy and the potential contribution of agricultural biotechnologies in developing countries. Biomass Bioenergy 34:1427-1439

Rubin EM (2008) Genomics of cellulosic biofuels. Nature 454:841-845

Ruijssenaars HJ, Hartmans S (2001) Plate screening methods for the detection of polysaccharase-producing microorganisms. Appl Microbiol Biotechnol 55:143-149

Schmidt CW (2007) Biodiesel: Cultivating alternative fuels. Environ Health Perspect 115:A86-A91

Schwarz WH, Grabnitz F, Staudenbauer WL (1986) Properties of a Clostridium thermocellum Endoglucanase Produced in *Escherichia coli*. Appl Environ Microbiol 51:1293-1299

Schwarz WH (2001) The cellulosome and cellulose degradation by anaerobic bacteria. Appl Microbiol Biotechnol 56:634-649

Shallom D, Shoham Y (2003) Microbial hemicellulases. Curr Opin Microbiol 6:219-228

Shoham Y, Lamed R, Bayer EA (1999) The cellulosome concept as an efficient microbial strategy for the degradation of insoluble polysaccharides. Trends Microbiol 7:275-281

Sims REH, Mabee W, Saddler JN, Taylor M (2010) An overview of second generation biofuel technologies. Bioresour Technol 101:1570-1580

Stephens E, Ross IL, Mussgnug JH, Wagner LD, Borowitzka MA, Posten C, Kruse O, Hankamer B (2010) Future prospects of microalgal biofuel production systems. Trends Plant Sci 15:554-564

Taherzadeh MJ, Karimi K (2008) Pretreatment of lignocellulosic wastes to improve ethanol and biogas production: a review. Int J Mol Sci 9:1621-1651

Tengerdy RP, Szakacs G (2003) Bioconversion of lignocellulose in solid substrate fermentation. Biochem Eng J 13:169-179

Timilsina GR, Shrestha A (2011) How much hope should we have for biofuels? Energy 36:2055-2069

Turner P, Mamo G, Karlsson EN (2007) Potential and utilization of thermophiles and thermostable enzymes in biorefining. Microbial Cell Factories doi:10.1186/1475-2859-6-9

VanFossen AL, Lewis DL, Nichols JD, Kelly RM (2008) Polysaccharide degradation and synthesis by extremely thermophilic anaerobes. Ann N Y Acad Sci 1125:322-337

Vidal B, Dien B, Ting K, Singh V (2011) Influence of feedstock particle size on lignocellulose conversion-a review. Appl Biochem Biotechnol 164: 1405-21.

Viikari L, Alapuranen M, Puranen T, Vehmaanperä J, Siika-Aho M (2007) Thermostable enzymes in lignocellulose hydrolysis. Adv Biochem Eng Biotechnol 108:121-145

Vrancken K, Van Mellaert L, Anne J (2010) Cloning and expression vectors for a Gram-positive host, *Streptomyces lividans*. Methods Mol Biol 668:97-107

Wang Y, Wang X, Tang R, Yu S, Zheng B, Feng Y (2010) A novel thermostable cellulase from *Fervidobacterium nodosum*. J Molec Catal B 66:294-301

Weber C, Farwick A, Benisch F, Brat D, Dietz H, Subtil T, Boles E (2010) Trends and challenges in the microbial production of lignocellulosic bioalcohol fuels. Appl Microbiol Biotechnol 87:1303-1315

Wongwilaiwalin S, Rattanachomsri U, Laothanachareon T, Eurwilaichitr L, Igarashi Y, Champreda V (2010) Analysis of a thermophilic lignocellulose degrading microbial consortium and multi-species lignocellulolytic enzyme system. Enzyme Microb Technol 47:283-290

Yang D, Weng H, Wang M, Xu W, Li Y, Yang H (2010a) Cloning and expression of a novel thermostable cellulase from newly isolated *Bacillus subtilis* strain I15. Mol Biol Rep 37:1923-1929

Yang SJ, Kataeva I, Wiegel J, Yin Y, Dam P, Xu Y, Westpheling J, Adams MW (2010b) Classification of '*Anaerocellum thermophilum*' strain DSM 6725 as *Caldicellulosiruptor bescii* sp. nov. Int J Syst Evol Microbiol 60:2011-2015

Zamost BL, Nielsen HK, Starnes RL (1991) Thermostable enzymes for industrial applications. J Ind Microbiol Biotechnol 8:71-81

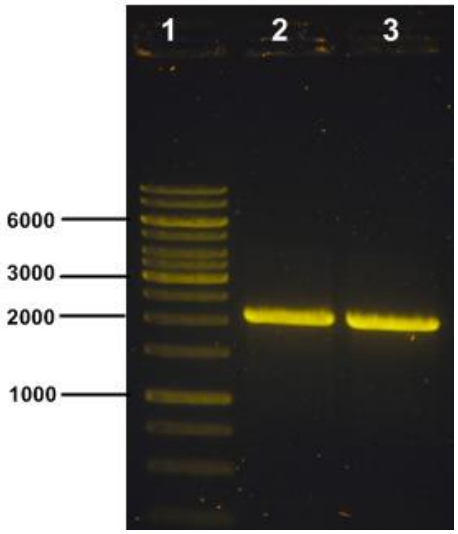
Zhang YH, Himmel ME, Mielenz JR (2006) Outlook for cellulase improvement: screening and selection strategies. Biotechnol Adv 24:452-481

Zhang YH (2008) Reviving the carbohydrate economy via multi-product lignocellulose biorefineries. J Ind Microbiol Biotechnol 35:367-375

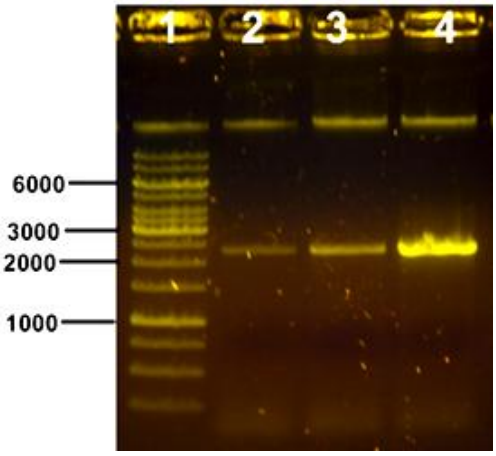
Zverlov V, Mahr S, Riedel K, Bronnenmeier K (1998) Properties and gene structure of a bifunctional cellulolytic enzyme (CelA) from the extreme thermophile '*Anaerocellum thermophilum*' with separate glycosyl hydrolase family 9 and 48 catalytic domains. Microbiology 144:457-465

Appendices

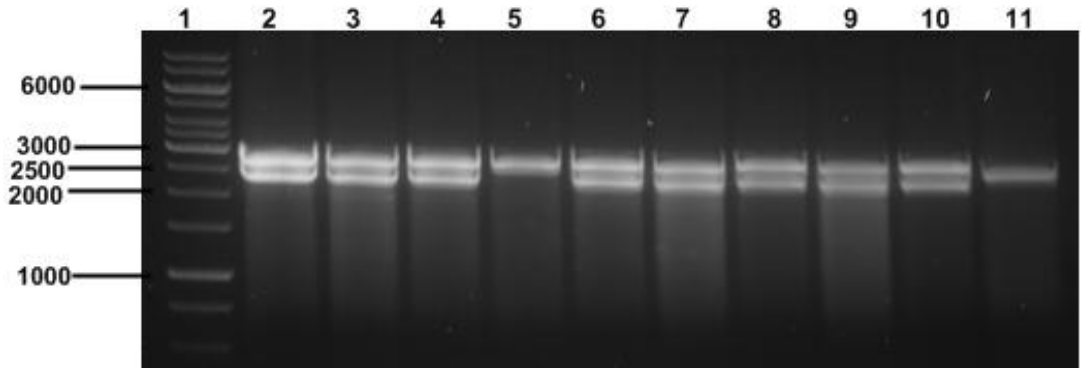
Appendix I: Agarose Gel electrophoresis images



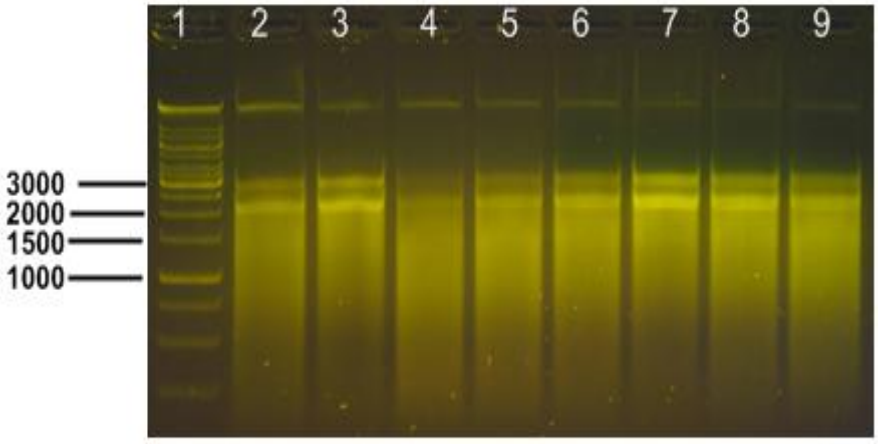
Appendix 1. PCR amplification of cellobiohydrolase gene (2127 bp) in lane 2 and 3. Lane 1 contain 1 kb DNA ladder from Fermantas Finland.



Appendix 2. PCR amplification of endoglucanase gene (2268 bp) in lane 2, 3 and 4. Lane 1 contains 1kb DNA ladder from Fermantas.



Appendix 3. Screening of transformant colonies harboring endoglucanase gene. Lane 1 contains DNA ladder (1kb) from Fermantas, lane 5 and 11 contains sample from control plates and remaining lanes contains plasmid DNA with endoglucanase gene (2268 bp) from screened colonies.



Appendix 4. Screening of transformant colonies containing exoglucanase gene. 1kb DNA ladder from Fermantas in lane 1 and remaining lane contains plasmid DNA with exoglucanase gene after digestion with *Xba*I and *Eco*RI endonucleases from screened colonies.

Appendix II: Experimental data observed

Appendix 5. Experimental data obtained from glucose standard curve assay.

Glucose conc (mg/ml)	OD-1	OD-2	Average	Standard devaiation
0,1	0,007	0,006	0,0065	0,000707107
0,25	0,084	0,095	0,0895	0,007778175
0,5	0,169	0,172	0,1705	0,00212132
0,75	0,271	0,27	0,2705	0,000707107
1	0,384	0,388	0,386	0,002828427
1,25	0,494	0,495	0,4945	0,000707107
1,5	0,638	0,607	0,6225	0,02192031
2	0,827	0,826	0,8265	0,000707107

Appendix 6. Values obtained during the optimum temperature experiment.

Temperature	OD-1	OD-2	Average	Glucose Conc.	Std Dev	Cmase activity [IU ml-1]
40	0,114	0,118	0,116	0,469939523	0,002828	0,348074074
50	0,127	0,126	0,1265	0,507292778	0,000707	0,375772428
60	0,132	0,131	0,1315	0,525080043	0,000707	0,38894818
65	0,136	0,134	0,135	0,537531128	0,001414	0,398171206
70	0,136	0,136	0,136	0,541088581	0	0,400806356
75	0,119	0,12	0,1195	0,482390608	0,000707	0,357326377
80	0,113	0,113	0,113	0,459267165	0	0,3401979
85	0,093	0,093	0,093	0,388118107	0	0,287494894
90	0,086	0,086	0,086	0,363215937	0	0,269048843

Appendix 7. Measurement results obtained during optimum pH determination.

pH	OD-1	OD-2	Average	Glucose conc.	Cmase activity [IU ml-1]	Std Dev
3	0,022	0,029	0,0255	0,147990039	0,109622251	0,00494975
4	0,126	0,125	0,1255	0,503735326	0,373137278	0,00070711
5	0,168	0,157	0,1625	0,635361081	0,470637838	0,00777817
6	0,134	0,14	0,137	0,544646033	0,403441506	0,00424264
7	0,114	0,116	0,115	0,46638207	0,3454682	0,00141421
8	0,11	0,113	0,1115	0,453930985	0,336245174	0,00212132
9	0,109	0,109	0,109	0,445037353	0,329657299	0
10	0,103	0,087	0,095	0,395233013	0,292765195	0,01131371

Appendix 8. Calculated enzyme activity during the thermostability test.

	At 70 degree		80 degree		90 degree	
Time (hr)	**CMase activity (IU ml-1)**	**Std Dev**	**Cmase activity (IU ml-1)**	**Std Dev**	**Cmcase (IU ml-1)**	**Std Dev**
0	0,420719652	0,00071	0,420719652	0,0007	0,420719652	0,0007
0,5	blank	blank	0,400948992	0	blank	blank
1	blank	blank	0,400948992	0	blank	blank
1,5	0,441808356	0,00071	0,400948992	0	0,049031238	0,0007
3	0,437854224	0	0,379860287	0	0,060893634	0
4,5	0,424673784	0	0,358771583	0	0,060893634	0
6	0,420719652	0,00071	0,358771583	0	0,062211678	0,0007
12	0,418083564	0,00071	0,258600237	0	0,05957559	0,0007
24	0,418083564	0,00071	0,159351522	0,0052	0,060893634	0

Appendix 9. Enzyme activity obtained with different substrate concentration.

CMC conc (mg/ml)	OD-1	OD-2	Average	glucose conc	Cmase activity [IU ml-1]
0,5	0,011	0,012	0,0115	0,09822064	0,07275603
1	0,026	0,028	0,027	0,15338078	0,113615395
2	0,05	0,051	0,0505	0,23701068	0,175563464
3	0,089	0,087	0,088	0,37046263	0,274416766
4	0,115	0,117	0,116	0,47010676	0,348227231
5	0,179	0,152	0,1655	0,64626335	0,478713589
6	0,185	0,188	0,1865	0,72099644	0,534071438
7	0,206	0,21	0,208	0,7975089	0,590747331
8	0,239	0,243	0,241	0,91494662	0,677738236
9	0,272	0,276	0,274	1,03238434	0,764729142
10	0,309	0,312	0,3105	1,16227758	0,860946356
11	0,33	0,333	0,3315	1,23701068	0,916304205

Appendix 10. Conversion of enzyme activity obtained from table A.5 into Lineweaver-Burk and Monod chart for determination of Vmax and Km value of endoglucanase enzyme.

Lineweaver-Burk and Monod Chart				
S	V	1/s	1/v	V
0,5	0,07275603	2	13,74456522	0,06863041
1	0,113615395	1	8,80162413	0,12898067
2	0,175563464	0,5	5,695945946	0,23018933
3	0,274416766	0,33333333	3,644092219	0,31172394
4	0,348227231	0,25	2,871688115	0,37881299
5	0,478713589	0,2	2,088931718	0,43498308
6	0,534071438	0,16666667	1,872408687	0,48269925
7	0,590747331	0,14285714	1,692771084	0,52373649
8	0,677738236	0,125	1,475495916	0,55940535
9	0,764729142	0,11111111	1,307652534	0,59069459
10	0,860946356	0,1	1,161512554	0,61836418
11	0,916304205	0,09090909	1,091340621	0,64300787